



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

행정학석사학위논문

지능정보정책에 관한 정부자료 분석

- 중앙행정기관 보도자료를 중심으로 -

2018년 8월

서울대학교 행정대학원

행정학과 행정학전공

최 한 별

지능정보정책에 관한 정부자료 분석

- 중앙행정기관 보도자료를 중심으로 -

지도교수 김 동 욱

이 논문을 행정학석사학위논문으로 제출함

2018년 3월

서울대학교 행정대학원

행정학과 행정학전공

최 한 별

최한별의 석사 학위논문을 인준함

2018년 6월

위 원 장 전 영 한 (인)

부위원장 엄 석 진 (인)

위 원 김 동 욱 (인)

국문초록

지능정보기술은 정보통신 분야에 한정되지 않고 다양한 분야에서 활용될 수 있는 범용기술로서의 특징과 사회·경제·문화 전반에 광범위한 파급력을 갖는다는 특징을 가지고 있다. 지능정보기술은 제조업, 전자상거래, 금융뿐만 아니라, 주거, 교통, 도시설계 등 사회 전반에 광범위한 영향을 주고 있으며, 앞으로도 다양한 기술 및 산업과 융합하여, 사회·경제·문화에 이전에 생각지도 못한 대대적인 사회변화를 일으킬 것으로 전망된다.

본 연구는 지능정보기술과 관련하여 정부 기관이 인식하고 있는 현재의 모습에 대한 분석이 필수적이라는 문제의식 하에 논의를 진행한다. 이에 중앙행정기관 지능정보정책의 전개 양상과 특징은 무엇인가라는 연구문제를 가지고, 지능정보정책에 관한 보도자료를 분석하였다. 방법론으로는 텍스트를 알고리즘에 의해 분석하는 텍스트 마이닝(text mining) 기법을 사용하여 분석하였다. 분석대상 자료는 2014년 1월 1일부터 2018년 6월 30일까지의 지능정보정책에 관한 중앙행정기관의 보도자료이다.

분석 결과로 첫째, 중앙행정기관의 지능정보정책 관련 보도자료에 나타난 주요 핵심어 흐름의 시기별, 기관별 활용 초점이 다름을 확인하였다. 둘째, 지능정보정책 관련 보도자료에 사용된 핵심어 단어분포의 집중도가 시간이 지날수록 커짐을 확인하였다. 셋째, 사회기술시스템 관점에서 지능정보정책 관련 보도자료는 산업·경제, 과학·기술적인 환경이 사회·제도적인 환경에서 설명되는 단어의 비중이 높음

을 확인하였다.

본 연구는 텍스트 형태의 비정형 데이터를 알고리즘을 이용하여 해석함으로써 정부가 수행하고 있는 정책에 관한 현재의 모습을 보다 객관적으로 분석할 수 있는 가능성을 제시한다는 의미가 있다. 기존의 정성적 방법에 의한 내용분석은 연구자의 주관적 관점에 따라 요약과 해석이 자의적일 수 있으며, 무엇보다 텍스트의 분량이 방대할 경우 적용이 어렵다. 텍스트 마이닝을 통한 정부 자료의 분석을 통해 이와 같은 한계를 보완할 수 있다.

중앙행정기관의 지능정보정책은 과학·기술, 산업·경제, 사회·제도 전 분야의 변화에 맞추어 각 분야가 긴밀하게 된 종합 정책을 지향하고 있다. 하지만, 아직까지 정책의 관심은 기술적 측면에서 경쟁력 확보와 산업적 활용 측면에 집중되어 있음을 확인하였다. 사회기술시스템 관점에서 지능정보기술 도입과 관련 정책의 안착이 성공적으로 이루어지기 위해서는 과학·기술의 도입과 산업·경제 측면에서의 활용뿐만 아니라, 사회 변화에 대응하는 교육 혁신, 일자리 변화 대응, 법제도 정비 및 윤리 정립과 같은 사회·제도적 주요 쟁점을 다루는 정책 마련이 필요하다.

주요어 : 지능정보정책, 지능정보기술, 사회기술시스템, 텍스트마이닝,
보도자료

학 번 : 2012-23756

목 차

| | |
|-----------------------------|----|
| 제 1 장 서론 | 1 |
| 제 1 절 연구의 목적과 필요성 | 1 |
| 제 2 절 연구의 범위 | 4 |
| 제 2 장 이론적 배경 및 연구 동향 | 6 |
| 제 1 절 사회기술시스템론에 대한 논의 | 6 |
| 제 2 절 지능정보정책에 대한 논의 | 10 |
| 제 3 절 연구방법에 대한 논의 | 14 |
| 제 3 장 연구문제 및 연구방법 | 17 |
| 제 1 절 연구문제 | 17 |
| 제 2 절 연구방법 | 20 |
| 1. 자료 수집 | 21 |
| 2. 주요 핵심어분석 | 23 |
| 3. 단어 분포의 집중도 분석 | 24 |
| 4. 사회기술시스템 척도 분석 | 25 |
| 제 4 장 분석 결과 | 27 |
| 제 1 절 자료의 분포 | 27 |
| 제 2 절 주요 핵심어 분석 | 31 |
| 1. 시기별 비교 | 31 |
| 2. 주요 기관별 비교 | 45 |
| 제 3 절 단어분포의 집중도 분석 | 51 |
| 1. 시기별 비교 | 51 |
| 2. 기관별 비교 | 52 |

| | |
|---------------------------|----|
| 제 4 절 사회기술시스템 척도 분석 | 54 |
| 1. 시기별 비교 | 54 |
| 2. 주요 기관별 비교 | 57 |
| 제 5 장 결론 | 62 |
| 제 1 절 연구결과 요약 | 62 |
| 제 2 절 연구의 함의와 제언 | 65 |
| 제 3 절 연구의 한계 | 66 |
| 참고문헌 | 68 |
| Abstract | 72 |

표 목차

| | |
|---|----|
| <표 1> 연구문제 | 19 |
| <표 2> 연결중심성 상위 30개 단어 (전체시기) | 32 |
| <표 3> 연결중심성 상위 30개 단어 (2014. 1.~2015. 12.) | 34 |
| <표 4> 연결중심성 상위 30개 단어 (2016. 1.~2016. 12.) | 37 |
| <표 5> 연결중심성 상위 30개 단어 (2017. 1.~2018. 6.) | 41 |
| <표 6> 연결중심성 상위 30개 단어 (주요기관) | 46 |
| <표 7> 지니계수를 활용한 단어분포의 집중도(시기별) | 51 |
| <표 8> 지니계수를 활용한 단어분포의 집중도(기관별) | 53 |
| <표 9> 사회기술시스템 척도 (시기별) | 55 |
| <표 10> 사회기술시스템 척도 (기관별) | 57 |
| <표 11> 사회기술시스템 척도 (과학기술정보통신부) | 60 |
| <표 12> 사회기술시스템 척도 (산업통상자원부) | 60 |
| <표 13> 사회기술시스템 척도 (국토교통부) | 61 |
| <표 14> 사회기술시스템 척도 (행정안전부) | 61 |

그림 목차

| | |
|---|----|
| <그림 1> 사회기술시스템의 구성요소 | 9 |
| <그림 2> 연구과정 요약 | 21 |
| <그림 3> 주요 키워드 및 관련 보도자료 분포 | 28 |
| <그림 4> 중앙행정기관별 보도자료 분포 현황 | 29 |
| <그림 5> 주요 키워드 및 관련 보도자료의 반기별 분포 | 29 |
| <그림 6> 핵심어 간 네트워크 지도 (전체시기) | 32 |
| <그림 7> 핵심어 간 네트워크 지도 ('14년 상반기) | 35 |
| <그림 8> 핵심어 간 네트워크 지도 ('14년 하반기) | 35 |
| <그림 9> 핵심어 간 네트워크 지도 ('15년 상반기) | 36 |
| <그림 10> 핵심어 간 네트워크 지도 ('15년 하반기) | 36 |
| <그림 11> 핵심어 간 네트워크 지도 ('16년 상반기) | 39 |
| <그림 12> 핵심어 간 네트워크 지도 ('16년 하반기) | 39 |
| <그림 13> 핵심어 간 네트워크 지도 ('17년 상반기) | 42 |
| <그림 14> 핵심어 간 네트워크 지도 ('17년 하반기) | 42 |
| <그림 15> 핵심어 간 네트워크 지도 ('18년 상반기) | 43 |
| <그림 16> 핵심어 간 네트워크 지도 (과학기술정보통신부) | 47 |
| <그림 17> 핵심어 간 네트워크 지도 (산업통상자원부) | 47 |
| <그림 18> 핵심어 간 네트워크 지도 (국토교통부) | 48 |
| <그림 19> 핵심어 간 네트워크 지도 (행정안전부) | 48 |

제 1 장 서론

제 1 절 연구의 목적과 필요성

지능정보기술은 인간의 인지, 학습, 추론과 같은 고차원적인 정보처리 활동을 기계를 통해 구현하는 기술을 의미한다(관계부처 합동, 2016; 정보통신기술진흥센터, 2016; 한국정보화진흥원, 2016). 대표적인 예로 인공지능(Artificial Intelligence), 사물인터넷(Internet of Things), 클라우드 컴퓨팅(Cloud Computing), 빅 데이터(Big Data), 스마트 시티(Smart City), 자율주행차(무인자동차), 사이버물리시스템(Cyber Physical System), 지능형 로봇(스마트 로봇) 등을 꼽을 수 있다. 지능정보기술은 정보통신 분야에 한정되지 않고 다양한 분야에서 활용될 수 있는 범용기술로서의 특징과 사회·경제·문화 전반에 광범위한 파급력을 갖는다는 특징을 가지고 있다. 지능정보기술은 적용 분야가 지속적으로 확대되고 있으며, 다양한 기술 및 산업과 융합하여 생산성과 효율성을 획기적으로 높이는 중심 역할을 하고 있다. 지능정보기술은 제조업, 전자상거래, 금융뿐만 아니라, 주거, 교통, 도시설계 등 사회 전반에 광범위한 영향을 주고 있으며, 앞으로도 다양한 기술 및 산업과 융합하여, 사회·경제·문화에 이전에 생각지도 못한 대대적인 사회변화를 일으킬 것으로 전망된다.

이렇게 사회 전 영역에서 활용될 수 있는 범용기술 특성을 가진 지능정보기술이 사회 전반에 혁신을 유발하고 광범위한 사회, 경제적 파급력을 가진 사회를 지능정보사회라고 할 수 있다(관계부처 합동, 2016; 성욱준·황성수, 2017; 한국정보화진흥원, 2016). 지능정보사회로의 변화는 과거 산업사회, 정보사회로 변화와 마찬가지로, 산업구

조, 고용구조뿐만 아니라, 교육, 문화, 복지 등 삶의 모습 전반의 변화를 초래할 것으로 보인다. 이렇게 지능정보기술로 인한 변화를 과거 증기기관, 전기 등이 사회에 대대적인 변화를 촉발한 수준으로 보는 관점에서 최근 ‘4차 산업혁명’으로 일컫기도 한다(Schwab, 2016).¹⁾ 대한민국에서는 2016년 3월 이세돌이 구글 딥마인드 社가 개발한 인공지능 바둑프로그램 알파고(AlphaGo)와의 바둑 대결 이후, 사회적으로 큰 관심이 일어났다.

지능정보사회로의 진입에 대비한 정부의 역할과 관련하여, 과거 경제성장 시대 수준으로 정부가 계획적으로 사회적 의제를 주도하고, 특정산업을 진흥하기는 어렵다. 그러나 정부가 민간의 역할이 극대화될 수 있도록 시장 환경을 개선하고, 인적자원 양성, 지능정보문화 확산 등의 사회 기반을 조성하는 역할은 여전히 중요하다. 지능정보 기술 관련 산업을 진흥하는 정책과 역기능에 대비하여 규제하는 정책을 균형 있게 추진하고, 여기에 공공부문이 선도적으로 지능정보기술을 도입해 민간 투자의 마중물 역할을 하는 등 지능정보사회를 실현하기 위한 정부의 정책적 역할은 여전히 필요한 상황이다.

한편, 기술의 활용에는 다양한 이해당사자가 관련되어 있다. 기술이 개발되고 도입되었을 때, 그 효과는 기술이 도입된 영역에만 영향을 미치는 것이 아니라, 사회 다양한 분야에 복합적으로 영향을 미치게 된다. 이에 따라 사회적 환경과 기술적 환경을 각각의 개별적으로만이 아니라 둘의 상호작용도 관리해야 기술을 둘러싼 복잡한 현상들을 이해할 수 있고, 기술을 효과적으로 활용할 수 있다. 따라서 기술의

1) 4차 산업혁명이라는 용어는 학술적으로 엄밀히 정의되는 용어는 아니지만, ‘제4차 산업혁명에 대응한 지능정보사회 중장기 종합대책’ (관계부처 합동, 2016), ‘혁신성장을 위한 사람 중심의 4차 산업혁명 대응계획’ (관계부처 합동, 2017)과 같은 정부 문건이나 연구 보고서에 자주 사용되고 있다.

도입과 활용 및 그 영향을 이해하기 위해서는 기술 그 자체에만 관심을 가지는 것만으로는 부족하며 보다 ‘포괄적인(holistic) 시선 또는 접근법’이 요구된다(어규철, 2011). 이렇게 사회와 기술은 서로 분리되어 존재할 수 없기에 상대편을 전제로 하는 사회기술시스템으로 존재한다고 파악하는 접근을 사회기술시스템(socio-technical system) 접근이라고 한다(Geels, 2002; 송위진, 2013). 지능정보기술은 정보통신 분야에 한정되지 않고 다양한 분야에서 활용될 수 있는 범용기술로서의 특성을 보유하고 있으며, 사회·경제·문화 전반에 광범위한 파급력을 가지기 때문에 이와 같은 접근이 더욱 필요하다.

본 연구는 중앙행정기관이 생산한 보도자료를 대상으로 지능정보정책의 전개 양상과 특징을 고찰함을 목적으로 한다. 본 연구는 보도자료가 정부의 정책 관심도와 정책 방향을 간접적으로 확인할 수 있는 유용한 분석의 대상임을 바탕으로 한다. 이에 지능정보기술 관련 키워드로 정부 보도자료를 수집하여 중앙행정기관의 지능정보 정책 동향을 파악하고, 사회기술시스템 관점에서 해석하였다.

제 2 절 연구의 범위

본 연구의 시간적 범위는 2014년 1월 1일부터 2018년 6월 30일까지로 한다. 최근의 지능정보기술에 대한 관심은 2016년 3월 프로바둑기사 이세돌과 구글 딥마인드社의 인공지능 바둑 프로그램인 알파고(AlphaGo)의 대국으로 인해 크게 촉발되었다. 본 연구의 시간적 대상인 2014년 1월 1일부터 2018년 6월 30일의 기간은 아직까지 지능정보기술에 대한 본격적인 사회적 관심이 미약한 시기(2014년~2015년)부터, 이세돌과 알파고의 바둑대결이 이후 2016년 3월 보다도 사회적 관심이 본격화되는 시기(2016년), 관련 범정부 종합대책이 수립된 이후, 지능정보기술에 대한 정책에 대한 논의가 본격화되는 단계(2017년~2018년)로 이어지고 있다.

본 연구의 대상은 지능정보정책에 관한 중앙행정기관의 보도자료로 하였다. 보도자료 외 지능정보 정책에 관한 범정부 종합대책도 존재하나, 범정부 종합대책 작성에는 작성에 관여하는 특정 부처의 비중이 강하고, 지능정보기술의 현재에 대한 객관적 인식보다는 나아가야 할 바람직한 방향에 대한 제시에 중점을 두고 있다는 점에서 분석 대상으로서 한계가 있다. 반면, 보도자료는 개별 중앙행정기관이 모두 배포할 뿐만 아니라, 중앙행정기관이 집행하는 정책의 반영이 즉각적이고, 양적으로도 풍부하여 연구대상으로 적합하다고 보았다. 이에 본 연구의 분석 대상은 전체 중앙행정기관의 보도자료로 하였다.

보도자료는 정책브리핑(www.korea.kr)에 지능정보 정책에 관한 주요 키워드를 입력하여 수집하였다. 검색 키워드는 지능정보사회에 대비한 정부 정책의 근간이 되는 보고서인 관계부처 합동(2016)과 지능정보사회에 대비한 정부 역할을 논의한 한국정보화진흥원(2016)에

서 지능정보기술의 주요 구성요소로 언급된 ‘인공지능’, ‘사물인터넷’, ‘클라우드 컴퓨팅’ (Cloud Computing), ‘빅데이터’ (Big Data), ‘스마트 시티’, ‘자율주행차’, ‘사이버물리시스템’ (CPS), ‘드론’, ‘지능형 로봇’ 을 꼽으면서 위 키워드를 통칭하는 ‘지능정보’ 와 ‘4차 산업혁명’ 을 함께 꼽았다.²⁾ 이를 통해 수집하여 분석 대상이 된 보도자료는 총 1,143건, 12,755개 문장이다. 검색어로 사용한 ‘4차 산업혁명’ 등 11가지 주요 키워드는 총 12,206건 수집하였다.

2) ‘클라우드 컴퓨팅’ 은 줄여서 ‘클라우드’ 로 많이 사용하므로, ‘클라우드’ 만으로 검색되는 결과도 포함하였다. 그 외에도 ‘스마트 시티’ 는 ‘스마트 도시’, ‘자율주행차’ 는 ‘스마트 카, 무인자동차, 커넥티드 카’, ‘드론’ 은 ‘무인기, 무인비행장치, 무인비행기’, ‘지능형 로봇’ 은 ‘지능 로봇, 스마트 로봇’ 등으로 검색한 결과를 포함하는 등 같은 의미를 가진 단어를 최대한 포함시켰다.

제 2 장 이론적 배경 및 연구 동향

제 1 절 사회기술시스템론에 대한 논의

사회기술시스템론은 사회와 기술은 서로 분리되어 존재할 수 없는 상대편을 전제로 하는 통합된 시스템으로 존재한다고 파악한다(송위진, 2012). 기술의 도입과 활용에는 다양한 당사자가 관련되어 있다. 기술은 단지 인프라, 기능, 관리 능력 등 기술과 직접 관련되는 요소만을 포함하는 것이 아니라 특정 기술과 별로 상관이 없어 보이는 사회규범 또는 문화적 가정과 같은 무형의 요소들도 포함한다(어규철, 2011). 사회기술시스템론에서는 사회과학자들이 독립변수로 강조해왔던 사회는 사회기술시스템으로 존재하고, 또 공학자들이 강조해왔던 기술도 사회기술시스템으로 존재한다고 파악하면서 사회와 기술을 통합적 관점에서 접근 한다(송위진, 2012). 사회기술시스템은 기술적 요소, 과학적 요소, 정책 요소, 사회문화적 요소, 사용자 및 시장 요소 등으로 구성된다(한재각 외, 2013; Geels, 2004).

사회적 환경과 기술적 환경을 각각의 개별적으로만이 아니라 둘의 상호작용도 관리해야 기술을 둘러싼 복잡한 현상들을 이해할 수 있고, 기술을 효과적으로 활용할 수 있다. 따라서 기술의 도입과 활용과 그 영향을 이해하기 위해서는 기술 그 자체에만 관심을 가지는 것만으로는 부족하고 보다 넓고 ‘포괄적인(holistic) 시선 또는 접근법’이 요구된다(어규철, 2011). 이 관점은 정보기술 도입과 활용에도 적용되는데, Sawyer et al.(2003)에 따르면 정보기술 활용의 목적을 달성하기 위해서는 정보 기술의 ‘기술적인 측면’ (the technological)과 ‘사회적인 측면’ (the social)을 동시에 고려해야

한다.

사회기술시스템론과 대비되는 개념으로 사회와 기술을 분리해서 양자 간에 인과관계를 설정하는 ‘기술결정론’과 ‘사회결정론’이 있다. ‘기술결정론’은 과학기술이 자체의 발전 논리에 따라 발전하고 그 효과에 의해 사회가 변화한다는 관점이고, ‘사회결정론’은 사회가 그 자체의 발전 논리가 있고 이에 따라 기술도 변화한다는 관점이다(송위진, 2012).

사회기술시스템론의 관점으로 사회제도 환경을 분석한 선행 연구로 성지은·송위진(2010)은 탈(脫)추격형 혁신의 효과적 전개를 위해 기술만이 아니라 시장과 제도가 함께 고려되는 ‘통합적 혁신정책(integrated innovation policy)’이 필요함을 보였다. 성지은 외(2012)는 지속가능한 교통·식품·가정 시스템으로의 전환을 시도하면서 과학기술 전문가뿐만 아니라 인문사회과학 전문가가 함께 참여하는 네덜란드의 사례를 분석하였다. 한재각 외(2013)는 사회기술시스템적인 시각에서 적정기술을 통한 국제개발협력 활동을 논의할 필요성을 제기하였다.

한편, 사회기술시스템론은 공공조직이 직면한 외부환경, 내부환경 그리고 주요한 이슈를 분석하는 프레임으로서 공공 영역에서 정보기술의 도입을 설명하는데 특히 중요한 이론으로 간주되고 있다(Reddick, 2012; 김선경, 2014). 정보기술은 특히 하드웨어와 소프트웨어의 단순한 합이 아니라, 기술을 이용하는 사람과 기술을 도입하는 사회문화적 맥락과 복잡하게 상호작용하기 때문에 사회기술시스템 접근이 더욱 필요하다고 할 수 있다(어규철, 2011). 이에 따라 사회기술시스템론을 배경으로 한 선행연구의 상당수가 정보기술 환경에 주목하고 있다.

정보기술 환경에 주목한 선행연구로 어규철(2011)은 개발도상국 ICT 인력 양성 프로그램 개발에 사회기술시스템 접근을 시도하였다. 이현규(2013)는 Bostrom&Heinen(1997)의 관점에 따라 분석의 틀을 구성하고, 미리 각 단어의 키워드 그룹을 만들면서 빈도수 상위 100개 단어를 사회·기술적 변수로 분류하였다. 김선경(2014)은 빅데이터를 기회와 효용은 많은 반면 매우 높은 위험성과 비용을 파생시키는 사회기술시스템으로 인식하고, 외부환경, 내부환경, 주요 이슈를 분석하였다. Orlikowski&Iacono(2001)은 광대역 통신과 모바일 통신 기술의 성공적으로 활용을 위한 사회기술적 관점에 주목하였다. Sawyer et al.(2003)은 사회기술적 관점에서 광대역 네트워크 및 모바일 기술에 대한 정책을 살펴보았다. Chen&Nath(2008)은 바람직한 모바일 작업환경 구축과 성과 극대화를 위해서는 사회적 그리고 기술적 시스템을 결합하여야 한다고 밝혔다. Lyytinen&Newman(2008)은 정보시스템의 변화 과정을 사회기술시스템의 맥락에서 살펴보았다. 이하영·엄석진(2018)은 한국 정부의 지식관리의 성과와 영향요인을 논하면서 기술발현을 촉진하는 조직적 맥락으로 혁신문화의 중요성을 보여주었다.

본 연구는 위와 같은 맥락에서 지능정보정책에 관한 정부 보도자료를 살펴본다. 사회기술시스템의 기본적 요소로 생산 영역, 기술의 사용 영역과 이에 영향을 미치는 제도 요소로 다루고 있는 Geels(2004)의 관점에 따라 지능정보정책을 바라볼 것이다. 특히, 관계부처 합동(2016)은 지능정보사회에 대비하여 기술, 산업, 사회 분야별 정책방향과 전략과제를 도출하였으며, 4차산업혁명위원회(2017)는 ‘사람 중심의 4차 산업혁명’의 청사진을 제시하면서, 앞으로의 추진 방향을 산업·경제, 사회·제도, 과학·기술의 세 가지

측면으로 정리하였다. 사회기술시스템의 맥락과 상통하면서도 정부의 공식적인 정책 방향이라는 점에서 의미가 있다. 이에 따라 산업·경제, 사회·제도, 과학·기술의 세 분류를 분석의 틀로 사용한다.

<그림 1> 사회기술시스템의 구성요소



본 연구는 정책을 설명하는 핵심어가 사회기술시스템 관점에서 어느 속성을 가지는 지에 대한 해석을 시도한다는 점에서 이현규 (2013)의 접근을 따른다. 하지만, 연구자의 사전적인 판단으로 각 단어를 특정 사회기술적 변수로 온전히 분류하기는 어렵다는 판단 하에 상호 정보량(mutual information)으로 단어의 속성을 파악 하였다. 또, 네트워크맥락에서 단어의 중요성을 파악한다는 판단 하에 연결중심성(degree centrality)를 사용하였다는 점에서 선행 연구와의 차별점이 있다.

제 2 절 지능정보정책에 대한 논의

지능정보기술과 그로 인한 사회변화의 중요성에도 불구하고, 이에 대한 이슈와 정책대응을 살펴본 연구는 많지 않은 실정이다. 특히, 지능정보사회와 지능정보기술에 관한 학문적 정의가 엄밀하게 이루어지지 않고 있어 관련 지식은 주로 정부 정책 자료와 관련 기관의 연구 보고서에서 얻게 되는 경우가 많다.

지능정보사회의 개념이 본격화되기 이전에는 스마트사회(smart society)라는 개념을 사용하기도 하였다. 스마트사회(smart society)는 “기존의 정보사회(information society)가 갖고 있던 물리적 제약(시·공간의 제약)에서 벗어나 항상 인터넷에 접속되어 있는 스마트 디바이스(always connected smart devices)를 통해 정보를 생산, 공유, 활용하는 사회”로 정의할 수 있다(정준화·김동욱, 2013). 즉, 스마트사회는 고정형 PC에서 이동형 스마트 디바이스로 전환과 같이 모바일 기반 이동형 서비스의 확산과 융합기술을 통해 인간의 감성과 창의성을 강조하는 것을 특징으로 한다(명승환 외, 2012). 정준화·김동욱(2013)은 스마트사회의 특징으로 이동성(mobility or ubiquitous)과 융합(convergence)을 꼽으면서, 스마트 디바이스와 클라우드 컴퓨팅을 통해 항상 인터넷에 접속하고, 위치정보가 적극적으로 이용되며, 다양한 디바이스에서 생산되는 정보가 빅데이터 형태로 수집·분석됨을 설명하였다. 지능정보사회는 이와 같은 스마트사회 논의의 맥락이 확장된 것이라고도 볼 수 있다.

기존의 스마트사회(smart society)라는 개념이 이동형 스마트 디바이스와 클라우드 컴퓨팅 및 빅 데이터 기술을 통해 실현가능해지는 것과 마찬가지로 지능정보사회의 도래도 이와 관련되는 지능정보기술

을 통해 가능하다고 할 수 있다. Schwab(2016)은 4차 산업혁명의 선도 기술을 물리(physical), 디지털(digital), 생물(biological)의 세 분야로 나누면서, 무인운송수단, 3D 프린팅, 첨단 로봇, 사물인터넷, 블록체인, 공유경제, 유전공학, 합성생물학, 그리고 바이오프린팅 등을 제시하였다.

관계부처 합동(2016)은 지능정보사회에 대비한 정부 정책의 근간이 되는 보고서라고 할 수 있는데, 4차 산업혁명의 동인인 지능정보기술의 구성요소로, 인공지능 기술(AI; Artificial Intelligence)과 데이터 활용기술(ICBM; IoT, Cloud computing, Big data, Mobile)을 꼽았다. 정보통신기술진흥센터(2016)은 4차 산업혁명에 활용될 것으로 예상되는 핵심기술로서 사물인터넷 (IoT; Internet of Things), 사이버물리시스템(CPS; Cyber-Physical System), 빅데이터 그리고 인공지능을 언급하였다. IoT, CPS, 빅데이터 그리고 인공지능에 의한 4차 산업혁명은 시스템의 지능성이 월등해지고(초지능성), 모든 사물과 시스템이 연결되며(초연결성), 향후 일어날 일에 대한 예측이 가능(예측가능성)할 것이라고 전망하였다. 한국정보화진흥원(2016)은 지능기술에 대해서 “사물이 인간의 개입 없이 판단이나 의사결정을 하도록 지원하거나, 또는 사물 스스로 상황인지(Contextually aware) 등 사고능력을 갖도록 하기 위해 사물을 지능화하는데 필요한 각종 요소 기술(IoT, AI, 로봇틱스, 빅데이터 등)을 모두 포함하는 개념”이라고 정의하였다. 이 보고서에서는 IoT, AI, 로봇틱스, 빅데이터와 같은 지능기술들이 급속히 발전하고 핵심기술로 주목받으면서 범용기술(General Purpose Technology)로 진화할 것이라고 예측하였다.

지능정보기술의 발전이 불러올 새로운 사회인 지능정보사회에 대해

서도 다양한 개념정의가 시도되고 있다. ‘4차 산업혁명’이라는 용어는 지능정보기술의 파급력을 과거 증기기관, 전기 등이 사회에 대대적인 변화를 촉발한 수준으로 보는 관점이다(Schwab, 2016). 국가전략과 주요 정책을 논의하기 위해 2017년 8월 대통령 직속기구로 출범한 4차산업혁명위원회의 명칭에도 사용되었다. 다만, ‘4차 산업혁명’이라는 용어는 2016년 세계경제포럼(World Economic Forum)에서 시작된 상업적 수사로, 기존 산업혁명과 구분되는 특징을 찾기도 어렵다는 의미에서 비판적으로 바라보는 시각(김동욱, 2018)과 위와 같은 변화가 산업구조뿐만 아니라, 국가·시장·사회 전 영역에 걸쳐 일어나는 영향력 있다는 시각(성욱준·황성수, 2017)에 따라 본 연구는 ‘4차 산업혁명’ 대신 ‘지능정보사회’라는 용어를 사용하였다.

관계부처 합동(2016)은 지능정보사회에 대하여 “고도화된 정보통신 기술 인프라(ICBM; IoT, Cloud, Big data, Mobile)를 통해 생성·수집·축적된 데이터와 인공 지능(AI)이 결합한 지능정보기술이 경제·사회·삶 모든 분야에 보편적으로 활용됨으로써 새로운 가치가 창출되고 발전하는 사회”라고 정의하였다.

한국정보화진흥원(2016)은 학계·언론계·기업·연구소 등 관련 전문가(로봇공학, 뇌 과학, 컴퓨터과학, 사회학) 20인을 선정하여 지능정보사회 개념 정의를 위해 인터뷰한 바 있는데, 전문가들은 “사물-인간 간 상호작용 극대화”, “지능기술기반 고도화 사회”, “사고능력 개선 문제해결제고 가치창출 사회”, “연장적 정보사회”라는 키워드를 중심으로 응답하였다.

한편, 지능정보기술로 인한 변화에 대한 전망은 사회적으로 긍정적인 전망과 부정적인 전망이 혼재하는데, 사회 전반에 기계가 인간을

대신하여 일을 수행함으로써 생산성 향상, 근로시간 감소, 건강수명 증가 등 경제·사회적 혜택이 고루 확대되는 긍정적인 전망이 있는가 하면, 자동화로 인한 단순·반복 업무의 일자리 수요가 감소하고, 고부가가치 업무의 인력수요가 증가하는 고용 구조 변화를 통해 사회적인 갈등을 초래할 것이라는 부정적인 전망도 있다. 그동안 대한민국은 국가적인 정보화 추진을 통해 세계적인 수준의 ICT 경쟁력을 갖추었다는 평가를 받고 있으나, 지능정보기술로 인한 경제·사회·문화의 대변화는 기술·산업 중심의 정보화를 넘어 교육·고용·복지 등 사회정책도 포괄한 전 사회적인 변화이므로 국가적으로 적절한 대비책 마련이 필수적인 상황이라고 볼 수 있다.

지능정보사회와 지능정보기술 정책과 관련한 주요 연구를 살펴보면 다음과 같다. 먼저, 성욱준·황성수(2017)는 주요 보고서 및 연구를 중심으로 미래에 도래할 지능정보시대의 모습과 이슈를 조망해보고 해외사례 및 한국 사례를 통해 정책적 대응 방향성을 모색해 보았다. 김병운(2016)은 인공지능 분야 주요국의 동향을 분석하고, 국가 과학기술·ICT 정책 제언을 하였다. 김윤명(2016)은 지능정보사회로 변화, 특히 인공지능 도입에 따른 윤리적 쟁점과 법제 정비에 관한 제언을 하였다. 황종성(2016)은 지능사회 등장에 따른 패러다임 변화와 정책적인 함의를 살펴보았다.

본 연구는 아직까지 지능정보정책에 관한 논의가 충분하지 않은 상황에서 지능정보정책 동향을 파악하기 위한 탐색적 연구의 성격을 갖는다. 정부 정책을 파악함에 있어 보다 효과적이고 객관적으로 파악하기 위해 텍스트 마이닝(text mining) 기법을 사용하고, 지능정보기술이 도입된 환경을 사회기술시스템 관점에서 해석하였다는 점에서 선행 연구와 차별점이 있다.

제 3 절 연구방법에 대한 논의

텍스트 마이닝(text mining)은 텍스트 형태의 비정형 데이터를 수학적 알고리즘과 컴퓨터 연산과정에 의해 분석하는 연구방법을 의미한다. 전통적으로 텍스트 데이터의 내용을 분석하는 방법으로는 연구자가 텍스트 데이터를 직접 읽어 세밀하게 이해한 후, 이론과 시각에 따라 해석하며, 이를 비판적으로 분석하는 질적 연구 혹은 정성적 연구방법(qualitative research method)을 주로 사용하였다(백영민, 2017). 하지만, 텍스트의 분량이 방대할 경우, 이와 같은 방법의 적용은 어렵고, 내용의 요약과 이론에 따른 해석도 텍스트를 직접 읽은 연구자의 주관적 관점에 따라 자의적일 수 있다는 문제가 있다(백영민, 2017; DiMaggio et al., 2013; Krippendorff, 2013).

이에 양적 연구 혹은 정량적 연구방법(quantitative research method)을 위와 같은 문제점을 보완하기 위한 대안적 방법으로 택할 수 있다. 텍스트를 수치형 데이터와 같이 간주하는 ‘텍스트 데이터 접근’(text-as-data approach)을 통하여, 분석 대상인 텍스트에 수학적 알고리즘과 컴퓨터 연산 과정을 적용시키고, 어떤 텍스트 구성요소가 분석 대상이 되는 텍스트를 주도적으로 설명하는지, 혹은 어떤 텍스트 구성요소가 분석 대상이 되는 텍스트의 의미를 예측하는데 효과를 발휘하는지를 정량화하는 방식이다(백영민, 2017; Gentzkow et al., 2017; Grimmer&Stewart, 2013). 이렇게 볼 때, 텍스트 마이닝은 대용량 텍스트 처리를 위해 비정형 데이터를 정량적 연구방법(quantitative research methods)을 사용하여 기존의 정성적 연구방법(qualitative research methods)의 한계를 극복하는 유용한 수단으로 사용될 수 있다.

비정형 텍스트를 수치형 데이터와 같이 접근하기 위해서는 분석에 방해가 되는 요소를 제거하는 사전 처리과정이 필요하다. 숫자표현 제거, 불용어 처리, 문장부호와 특수문자 제거 등의 기술적 처리 과정이다. ‘무인기-드론’ 처럼 같은 의미로 사용된 단어가 다른 단어로 인식되지 않게 동의어처리를 하는 것도 중요한 과정이다. 한국어의 경우, 단어에 조사(助詞)가 붙어 있고, 동사의 활용도 다양하여 띄어쓰기에 의한 단어 구분이 어려운 등 언어 특성이 영문 텍스트의 사전 처리 과정과 다르다. 따라서 영문 텍스트를 분석하는 방법을 그대로 적용시킬 수 없다. 이렇게 사전처리 과정을 통해 텍스트 데이터가 명사 단위로 정제가 되면, 이를 이용하여 본격적인 분석(빈도표 계산, 상관관계 분석, 토픽 모델링 등)을 할 수 있다.

문서와 단어로 구성된 행렬(문서×단어 행렬, document-term matrix, DTM)은 가로줄과 세로줄로 구성된 행렬이라는 점에서 다른 사회과학 데이터와 본질적으로 유사한 구조를 갖는다(백영민, 2017). 이때, 특정단어 $word_i$ 가 등장했을 때, 다른 특정단어 $word_j$ 가 동시에 등장하는 연관관계를 수치화한 계산도 가능할 것이다. 즉, 텍스트 내에서 동시에 출현하는 단어 중 통계적으로 유의미한 관계를 보이는 단어들을 추출함으로써 텍스트의 의미 구조 분석이 가능하다.

텍스트 마이닝 기법을 사용하여, 사회 담론이나 정부 정책을 분석한 주요 연구는 다음과 같다. 정준화·김동욱(2013)은 개인정보 위험을 다룬 신문기사에 대해 텍스트네트워크분석을 시행하여 개인정보 위험 대상과 방식에서 스마트사회가 기존의 정보사회와 어떠한 차이점이 있는지 분석하였다. 텍스트네트워크분석을 통해 분석대상의 양적인 차이와 질적인 차이를 측정하였다. 이은미 외(2016)는 행정자치부의 보도자료를 텍스트 마이닝을 통해 분석함으로써 개인정보보호 제도의 내생적 변화를 살펴보았다. 통계적인 방법을 적용하기 어려운 제도 현상 분석에 주로 사용되는 서술적인 분석(narrative analysis)의 객

관성 확보를 위한 대안적 방안을 모색하였다. 심주영(2017)은 용산 미군기지 공원화 과정에서 생산된 대중담론을 신문기사 자료의 텍스트 마이닝을 통해 해석하였다. 용산 미군기지 공원화 과정을 구간으로 구분하여, 토픽분석과 네트워크 분석을 시행하였고, 이 결과를 도시 담론 관점에서 해석하였다. 서병조·신선영(2017)은 지능정보사회 대비를 위한 플랫폼 정부의 연구동향을 토픽 모델링을 통해 해석하였다. 플랫폼 정부가 활발히 논의된 시기와 중심 키워드 변화를 도출함으로서 국내 연구자들의 관심 방향 분석을 시도하였다.

선행연구를 통해 살펴본 것처럼 텍스트 마이닝 기법은 방법론적으로 다양한 연구에 적용될 수 있다. 본 연구는 위와 같은 맥락에서 지능정보정책 등 정부의 특정 분야 정책의 동향 파악에 텍스트 마이닝 기법을 사용한다. 본 연구는 내용 분석 등 기존의 정성적 연구방법으로 분석이 어려운 문제를 컴퓨터 연산을 이용해 해결하는 방법론인 텍스트 마이닝을 통해 분석하고자 한다.

제 3 장 연구문제 및 연구방법

제 1 절 연구문제

지능정보정책에 관한 선행연구는 정부의 지능정보 정책에 관한 주요 내용과 논점을 정리하고, 정책적 방향성을 짚어주고 있다. 그럼에도 불구하고, 아직까지 관련 연구가 초기적인 수준에 있으며, 양적으로도 관련 논의를 이어가는데 충분하지 않다. 정부의 정책 방향과 제언을 제시하는 방향으로 논의가 이루어지고 있음에도 정부 기관이 인식하고 있는 지능정보사회의 현재에 대한 분석은 부족하다. 또, 사회적 환경과 기술적 환경을 동시에 고려할 필요가 있다. 마지막으로, 정성적 연구방법(qualitative research method)에 입각해 연구자의 주관과 평가에 크게 좌우 받을 가능성이 있다. 본 연구는 위와 같은 한계를 보완하고, 지능정보기술 관련 정부 기관이 인식하고 있는 현재의 모습에 대한 분석이 필요하다는 문제의식 하에 논의를 진행한다. 지능정보정책에 대한 정부의 인식 현황은 어떠한 지에 대해 탐색적 연구한다. 이에 본 연구의 연구문제는 다음과 같다.

연구문제>중앙행정기관 보도자료에 나타난 지능정보정책의 전개 양상과 특징은 무엇인가?

연구문제를 구체화하기 위한 세부 과제는 다음과 같다.

첫째, 지능정보정책의 전반적인 현황을 분석한다. 이와 관련하여 본 연구는 시간의 흐름에 따른 지능정보정책에 관한 논의가 다양화·구체화됨을 볼 수 있을 것으로 예상하였다.

둘째, 시기별, 주요 중앙행정기관별 지능정보정책이 다루는 핵심단어의 분산된 정도를 분석한다. 2014년 1월부터 2018년 6월까지를 대상 기간으로 하며, 지능정보정책에 관한 보도자료의 상당수를 차지하는 과학기술정보통신부, 산업통상자원부, 국토교통부, 행정안전부의 4개 기관을 분석한다. 이와 관련하여 시기별로는 시간이 지날수록 중앙행정기관이 다루는 지능정보정책에 관한 주제가 구체화되기 때문에 단어 집중도는 높아질 것이라고 예상하였다. 한편, 주요 중앙행정기관별로는 전자정부와 행정혁신을 다루는 기관(행정안전부)이 다루는 주제가 가장 다양하여 단어 집중도가 가장 낮고, 과학기술정책을 총괄하는 기관(과학기술정보통신부)과 상업·공업에 관한 정책을 다루는 기관(산업통상자원부)이 그 다음일 것으로 예상하였다. 한편, 국토·교통 분야의 정책을 다루는 기관(국토교통부)이 다루는 주제는 특정 산업의 구체적인 모습을 보여줄 것이기 때문에 단어 집중도가 가장 높을 것으로 예상하였다.

셋째, 시기별, 주요 중앙행정기관별 지능정보정책을 사회기술시스템 관점에서 어떻게 해석할 수 있는 지 확인한다. 이와 관련하여 본 연구는 시기별로는 지능정보정책에 관한 논의는 초기에 주로 과학·기술 측면에서 논의되다가 기술의 활용이 본격화되면서 산업·경제 측면이 부각되며, 기술의 활용이 안정화되면서 점차 사회·제도 측면 설명이 많아질 것으로 예상하였다. 한편, 주요 중앙행정기관별로는 기관의 성격에 따라 과학기술정책을 총괄하는 기관(과학기술정보통신부)은 과학·기술 측면의 설명이 많을 것이며, 상업·공업에 관한 정책을 다루는 기관(산업통상자원부) 및 국토·교통 분야의 정책을 다루는 기관(국토교통부)은 산업·경제 측면의 설명이 많을 것으로 예상하였다. 한편, 전자정부와 행정의 능률·혁신을 다루는 기관(행정안

전부)은 상대적으로 사회·제도 측면의 설명이 많을 것으로 예상하였다. 이를 정리하면, <표 1>과 같다.

<표 1> 연구문제

| <연구문제> 중앙행정기관 보도자료에 나타난 지능정보정책의 전개 양상과 특징은 무엇인가? | |
|---|---|
| 전반적인 현황 | <1> 시간의 흐름에 따라 지능정보정책에 관한 논의의 변화를 살펴본다. |
| | <2> 주요 중앙행정기관별(과학기술정보통신부, 산업통상자원부, 국토교통부, 행정안전부) 지능정보정책에 관한 논의를 비교한다. |
| 핵심단어의 분산정도 | <3> 지능정보정책에 관한 단어 집중도의 변화를 시간의 흐름에 따라 살펴본다. |
| | <4> 지능정보정책에 관한 단어 집중도를 주요 중앙행정기관별(과학기술정보통신부, 산업통상자원부, 국토교통부, 행정안전부)로 비교한다. |
| 사회기술시스템 관점에서의 설명 | <5> 시간에 따른 지능정보정책의 변화를 과학·기술, 산업·경제, 사회·제도의 맥락에서 설명한다. |
| | <6> 주요 중앙행정기관별(과학기술정보통신부, 산업통상자원부, 국토교통부, 행정안전부) 지능정보정책을 과학·기술, 산업·경제, 사회·제도 맥락에서 비교한다. |

제 2 절 연구 방법

본 연구는 지능정보정책에 관한 이론과 선행 연구가 충분하지 않은 현실에서 변수 간의 인과관계를 살펴보는 대신에 정책의 현황과 전개 양상의 특징을 살펴보는 탐색적(exploratory) 연구의 성격을 갖는다.

방법론으로는 정부 보도자료에 담긴 지능정보 정책에 관한 의미를 객관적으로 분석하고, 자의적 해석의 오류를 피하기 위해 텍스트 마이닝(text mining) 기법을 사용하였다. 연구과정은 다음과 같다.

첫째, 보도자료의 문장을 컴퓨터 연산을 통해 분석할 수 있도록 명사 단위로 정리하는 전 처리 과정을 거친 후, 자료의 분포를 확인하는 기초 작업을 진행한다.

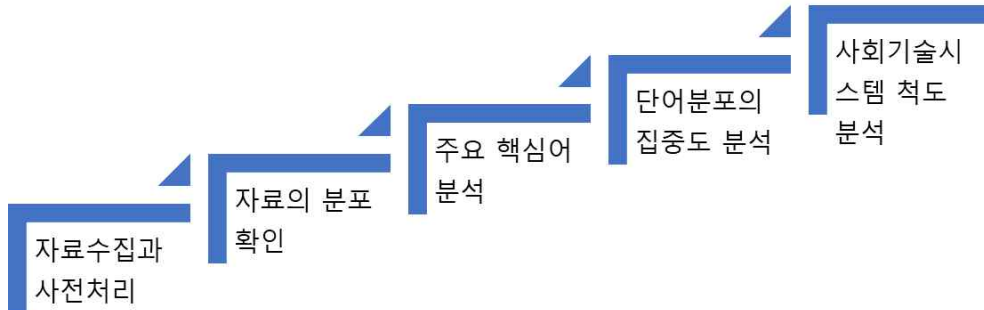
둘째, 지능정보정책의 전반적인 현황을 확인하기 위해 주요 명사 간의 연결에 기반 한 핵심어를 도출하고, 이를 해석하였다.

셋째, 핵심단어의 분산정도를 확인하기 위해 소득분포의 불평등도를 측정하는 지니 계수를 단어의 분포에 적용한 ‘단어분포의 집중도’를 도출하고, 이를 해석하였다.

넷째, 지능정보정책을 사회기술시스템 관점에서 해석하기 위해 핵심 단어의 연결중심성(degree centrality)과 핵심단어와 사회기술시스템 구성요소의 상호정보량(mutual information)을 이용하여 ‘사회기술시스템 척도’를 도출하고, 이를 해석하였다.

본 연구의 연결중심성 도출, 지니계수 도출 및 시각화는 Cyram(2018)의 Netminer 4.4.0.b 패키지를 이용하였다. 해당 패키지는 한글 형태소 분석기로 Seunjeon 1.5.0을 활용한다. 전체적인 연구과정은 <그림 2>와 같다.

<그림 2> 연구과정 요약



1. 자료 수집

정책브리핑(www.korea.kr)에 지능정보 정책에 관한 주요 키워드를 입력하여 수집하였다. 검색어는 지능정보사회에 대비한 정부 정책의 근간이 되는 보고서인 관계부처 합동(2016)과 지능정보사회에 대비한 정부 역할을 논의한 한국정보화진흥원(2016)에서 지능정보기술의 주요 구성요소로 언급된 ‘인공지능’, ‘사물인터넷’, ‘클라우드 컴퓨팅’ (Cloud computing), ‘빅데이터’ (Big data), ‘스마트 시티’, ‘자율주행차’, ‘사이버물리시스템’ (CPS), ‘드론’, ‘지능형 로봇’을 꼽으면서 위 키워드를 통칭하는 ‘지능정보’와 ‘4차 산업혁명’을 함께 꼽았다.³⁾ 위 문건이 정부의 정책을 수립하는데 기초적인 자료가 되거나 혹은 공식화된 정책방향이라는 점에서 정부의 지능정보정책을 읽을 때 활용할 수 있는 의미 있는 키워드라고

3) ‘클라우드 컴퓨팅’은 줄여서 ‘클라우드’로 많이 사용하므로, ‘클라우드’만으로 검색되는 결과도 포함하였다. ‘스마트 시티’는 ‘스마트 도시’, ‘자율주행차’는 ‘스마트 카, 무인자동차, 커넥티드 카’, ‘드론’은 ‘무인기, 무인비행장치, 무인비행기’, ‘지능형 로봇’은 ‘지능 로봇, 스마트 로봇’ 등으로 검색한 결과를 포함하는 등 같은 의미를 가진 단어를 최대한 포함시켰다.

보았다. 또한, 보도자료의 본문만을 수집하였으며, ‘붙임’ 등 참고 자료는 제외하였다. 이와 같이 수집하여 분석 대상이 된 보도자료는 총 1,143건, 12,755문장이다.

수집된 자료는 주로 *.hwp 혹은 *.pdf 형식으로 되어있으나, 본 분석에서는 글자 크기나 폰트 등 다른 정보는 필요하지 않고 텍스트 문자열 집합만이 의미 있으므로 분석의 편의를 위해 *.txt 형식으로 변환시킨다. 이후, 조사(助詞) 처리, 불용어 처리, 특수문자 제거 등의 사전처리 과정의 자료 정제(text refining)를 거친다. ‘스마트 도시’와 ‘스마트 시티’와 같이 같은 의미를 지닌 용어는 동의어 처리를 하고, ‘4차 산업혁명’과 같이 작은 단어가 여러 개 모여서 하나의 단어가 된 복합어는 한 단어로 인식할 수 있게 정의어 처리도 진행한다. 사전처리 과정을 거친 자료는 문서×단어 행렬(document-term matrix, DTM)로 변환하여 텍스트를 수치형 자료처럼 접근할 수 있도록 정리한다.

이렇게 수집한 텍스트의 각 시기별로 분류를 통해 지능정보정책에 관한 보도자료의 시기별 분포를 확인하고, 수집한 텍스트의 각 중앙행정기관별 분류를 통해 지능정보기술 정책에 관한 중앙행정기관별 분포를 확인한다. 또한, 수집된 텍스트의 각 키워드(인공지능, 사물인터넷, 클라우드 컴퓨팅, 빅데이터, 스마트 시티, 자율주행차, 사이버물리시스템, 드론, 지능형 로봇, 지능정보, 4차 산업혁명)별 분포를 확인할 수 있다.

2. 주요 핵심어분석

수집된 텍스트 데이터에 대한 기술통계분석은 일반적인 사회과학 수치 데이터에 대한 기술통계와 유사한 방식으로 진행할 수 있다.

상당수의 연구는 핵심어 분석을 시도할 때, 단어 빈도에 근거하여 제시한다(박주섭·홍순구, 2016). 단어 빈도를 통해 다수 등장한 단어가 해당 문서를 보다 잘 설명한다는 관점이다. 하지만, 본 연구의 경우, 보도자료의 특성상 배포한 중앙행정기관과 실무부서의 명칭, 특정인물의 이름, ‘이하’, ‘붙임’ 등과 같은 분석에 크게 의미가 있지 않은 단어의 빈도가 지나치게 높아지는 문제가 있다. 이에 따라 본 연구는 단순 빈도보다는 단어에 연결된 링크가 많을수록 문서를 설명하는 대표성 있는 단어라고 판단하여 연결중심성(degree centrality)⁴⁾ 개념을 이용하였다.

본 연구에서는 각 시기 및 주요 기관별 명사를 노드로 하는 의미연결망(window size=3)에서 두 단어가 2회 이상 함께 등장한 경우를 추출하였다. 즉, 분석 대상 문장을 기준으로 문장의 앞 뒤 세 문장까지를 한 범주로 하여, 동시에 두 번 이상 출현하는 단어를 연결하였으며, 다른 단어와 많이 연결된 단어를 핵심어로 본 것이다.

이렇게 시기별, 기관별로 정리한 텍스트 집단에 포함된 단어 간 링크를 확인하여 중심성이 높은 단어순으로 정렬할 수 있다. 이를 통해 텍스트 집단별로 어떤 단어가 해당 텍스트 집단을 잘 설명하는 지 핵심어를 확인할 수 있고, 추출된 주요 단어를 중심으로 핵심 내용 파악이 가능하다. 이를 통해 중앙행정기관의 지능정보정책의 시기별·주요 행정기관별 전반적인 현황의 분석이 가능하다.

4) 연결중심성(degree centrality)은 연결된 노드의 개수로 네트워크의 중심도를 나타내며, 다음과 같이 계산된다.

$$Cd = \frac{\sum [weight\ of\ incident\ links]}{nodes - 1}$$

3. 단어 분포의 집중도 분석

각 시기 및 주요 기관별로 사용된 단어 분포의 다양성을 확인하기 위해서 인구분포와 소득분포와의 관계를 소득분배의 불평등도에 적용할 수 있다. 관련 선행연구로 이창길(2010)은 노무현 정부와 이명박 정부의 임기 초기 대통령 연설문을 분석하면서 두 정부의 가치지향의 불균형성을 측정하는 지표로 지니계수를 이용하였다. 이은미(2018)는 지니계수와 로렌즈곡선을 사용하여, 민원 데이터 분포의 다양성을 살펴보았다.

본 분석에서는 지니 계수(Gini coefficient)⁵⁾를 단어의 종류와 단어와 연결된 링크의 숫자와의 관계에 적용하여 단어분포의 집중 정도를 확인하였다. 소득불평등도와 비교하여 일종의 ‘단어불평등도’를 측정하였고, 이를 통해 단어 분포의 집중정도를 확인한 것이다. 이를 통해 중앙행정기관의 지능정보정책의 시기별·주요 행정기관별 다루는 주제의 다양성의 정도를 확인할 수 있다.

5) 소득순위 백분위별 그룹의 누적소득비중을 나타낸 그래프로서 소득불균등 정도를 나타내는 곡선을 로렌즈곡선이라 한다(김영산·왕규호, 2013). 소득 분포가 균등할수록 직선에 가까운 그래프가 그려진다. 지니계수는 로렌즈 곡선이 그려진 백분위별 소득순위-누적소득비중 그래프에서 완전균등선(45도선) 아래의 면적 대비 완전균등선과 로렌즈곡선으로 둘러싸인 부분의 면적의 비율로서 소득이 완전히 균등한 경우에는 0이며, 이론적으로 가능한 최대값은 1이다(김영산·왕규호, 2013). 소득분포가 균등할수록 0에 가까운 값을 나타낸다.

4. 사회기술시스템 척도 분석

4차산업혁명위원회(2017)는 ‘사람 중심의 4차 산업혁명’의 청사진을 제시하면서, 앞으로의 추진 방향을 산업·경제, 사회·제도, 과학·기술의 세 가지 측면으로 정리하였다. 본 연구는 중앙행정기관의 지능정보정책을 위 관점에서 어느 정도 설명하고 있는지 해석하고, 각 시기와 주요 기관별 지능정보정책이 사회기술시스템 관점에서 어떠한 환경에 있는 지 파악을 진행한다. 이를 위해 단어의 연결중심성(degree centrality)과 상호정보량(mutual information)을 활용하여 ‘사회기술시스템 척도’를 도출하였다.

관련 선행연구로 사회기술적 접근 방식을 통해 모바일 환경을 분석한 이현규(2013)는 신문기사에 등장한 단어를 기술적 환경과 사회적 환경으로 이분하고, 빈도수 기준 상위 100개 중 각각 기술적 환경과 사회적 환경이 속한 숫자를 구했다. 연구자가 각 단어가 어느 속성을 갖는지 분류를 하고, 그 개수를 통해 환경을 분석한 것이다. 본 연구는 Geels(2004)의 관점을 적용하면서도 관계부처 합동(2016)과 4차산업혁명위원회(2017)에서 제시된 추진 방향을 참고해, 각 명사의 성격의 분류를 ① 과학·기술, ② 산업·경제, ③ 사회·제도의 속성 등 세 가지로 분류한다. 단, 선행연구와 차별되는 점은 각 단어의 속성을 사전에 판단하지 않고, 각 단어의 연관성 정도를 측정하는 상호정보량(mutual information)⁶⁾을 활용한다는 점이다. 단어의 동시출

6) 상호정보량은 두 사건 와 가 동시출현할 확률과 각각 독립적으로 출현할 확률 사이의 비율을 의미하며, 다음과 같다.

$$MI(x,y) = \log_2 \frac{P(x,y)}{P(x) \times P(y)}$$

상호 정보량은 두 확률이 완전히 독립적인 경우 0이 되고 의존관계가 높을수록 높은 값을 가진다(정영미·이재운, 1998).

한편, 출현빈도에 따라 최대값이 달라지는 현상을 보완하기 위해 다음과 같이 결합정보량 대비 상호정보량의 상대적 크기를 도출하여 상호정보량의 크기를 정규화 할 수 있으며 아래와 같다(이재운, 2003).

현빈도를 이용하여 도출하는 상호정보량은 단어 간 유사도 산출을 비롯하여 다양한 분야에서 연관성 척도로 사용된다(이재운, 2003).

상호정보량 도출은 2018년 7월 1일 기준 421,653개의 문서를 가져 상당한 정보를 보유하고 있는 위키피디아(wikipedia.org) 검색 정보를 단어사전으로 하여 도출하였다. 본 연구는 연결중심성이 0.07 이상인 단어의 연결중심성을 가중치로 위키피디아 검색 정보량을 활용하여 도출한 상호정보량과 과학·기술, 산업·경제, 사회·제도 각 속성을 곱하여 도출한다. 이때, 도출된 값은 상호비교를 위해 각 속성의 값의 합이 1이 되도록 표준화하였으며, 이 값을 ‘사회기술시스템 척도’로 본다.

이 방법의 선행연구와 비교하여 보완 점 및 차별 점은 다음과 같다. 먼저 각 단어의 속성을 연구자가 사전판단하지 않고, 방대한 문서사전 정보를 이용한다. 둘째, 각 단어가 어느 한 속성에 온전히 속한다고 가정하지 않는다. 셋째, 연결중심성을 가중치로 하여 각 단어의 중요도를 반영한다.

이렇게 분류된 단어를 연결중심성과 상호정보량을 활용하여 도출된 값을 통해 지능정보정책에 관한 주요 정책이 사회기술시스템론 관점에서 균형 있는 접근이 이루어져 있는 지 참고가 가능하도록 한다.

$$RM I_j(x, y) = \frac{MI(x, y)}{-\log_2 p(x, y)}$$

본 연구는 출현빈도에 따라 최대값이 달라지는 현상을 보완한 위 정규화한 상호정보량을 활용하였다.

구체적으로 위키피디아에 각 단어를 검색한 문서의 개수와 각각 ‘과학 기술’, ‘산업 경제’, ‘사회 제도’를 검색한 문서의 개수, 그리고 각 단어와 ‘과학 기술’, ‘산업 경제’, ‘사회 제도’를 동시에 검색한 문서의 개수를 활용하여 상호정보량을 구하는 방식이다.

제 4 장 분석 결과

제 1 절 자료의 분포

분석이 되는 보도자료는 총 1,143건의 12,755개 문장이며, 검색어로 사용한 ‘4차 산업혁명’ 등 11가지 주요 키워드는 총 12,206번 등장한다. 주요 키워드의 빈도와 키워드별로 수집된 보도자료의 빈도는 <그림 3>과 같다. 가장 많이 등장한 키워드는 ‘빅데이터’ (2,361건)이며, ‘드론’ (2,013건), ‘4차 산업혁명’ (1,718건), ‘클라우드 컴퓨팅’ (1,666건) 순으로 빈번하게 나타났다. 키워드별로 수집된 보도자료는 ‘4차 산업혁명’ (331건), ‘빅데이터’ (248건), ‘드론’ (151건), ‘스마트시티’ (98건) 순으로 많이 나타났다.

특이할만한 점은 다수의 연구에서 지능정보기술의 핵심으로 꼽고 있는 ‘사이버 물리 시스템’ 으로 검색되는 보도자료가 한 건도 존재하지 않는다는 것이다. 또, 최근 사회적으로 ‘4차 산업혁명’ 이라는 단어를 많이 사용하면서 ‘4차 산업혁명’ 으로 검색되는 보도자료 (331건)가 ‘지능정보’ 로 검색되는 보도자료(30건)보다 10배 이상 많다.

<그림 3> 주요 키워드 및 관련 보도자료 분포

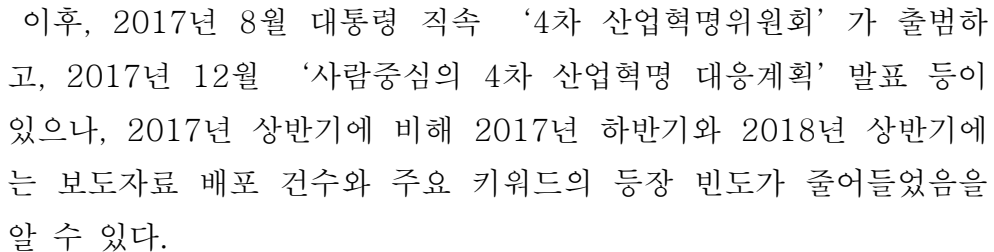
(단위 : 건)



위의 보도자료는 총 37개 중앙행정기관에서 배포한 것이며, 현황은 <그림 4>와 같다. ‘과학기술정보통신부’ (271건), ‘국토교통부’ (254건), ‘행정안전부’ (122건), ‘산업통상자원부’ (112건) 순으로 많다.⁷⁾ 특히, 위 네 개의 중앙행정기관에서 배포한 보도자료가 전체 보도자료의 66.4%를 차지하고 있어 지능정보정책에 관한 보도자료는 특정 부처에 집중되어 배포되고 있음을 알 수 있다.

7) 조직 명칭이 변경되거나 통합된 경우, 현재 존재하는 행정기관 명칭을 기준으로 정리하였다. 예를 들어 미래창조과학부의 보도자료는 과학기술정보통신부, 국민안전처의 보도자료는 행정안전부 통계에 포함된다.

(단위 : 건)



<그림 5> 주요 키워드 및 관련 보도자료의 반기별 분포

(단위 : 건)



제 2 절 주요 핵심어 분석

1. 시기별 비교

지능정보정책 관련 키워드를 통해 수집된 전체 보도자료는 1,143건으로 12,755개의 문장을 분석하였다. 각 시기 및 주요 기관별 명사를 노드로 하는 의미연결망(window size=3)을 도출하였으며, 두 단어가 2회 이상 함께 등장한 경우의 연결중심성을 상위 30개의 단어를 <표 2>와 같이 구하였다. 한편, 핵심어 간 연결망을 시각화하면 <그림 6>과 같다. 연결 횟수가 100회 이상이 되는 단어들로 구성되는 연결망을 시각화하였는데, 모든 연결 횟수를 반영하기에는 대상 단어가 5,950개로 과도하게 많기 때문에 육안으로 식별이 가능한 시각화를 위해 연결횟수가 일정 수준 이상인 핵심어의 네트워크로 제한하였다.⁸⁾

보도자료에서 연결중심성이 높게 나타난 단어는 ‘기술, 이용, 정보, 산업, 개발, 드론, 서비스, 기업’ 순으로 나타났다. 주요 단어를 통해 살펴보면, 중앙행정기관의 보도자료에서 지능정보 정책은 주로 ‘기술 개발과 이용에 대한 정보’와 ‘산업·기업을 지원’ 하는 측면에 초점이 있다고 볼 수 있다.

8) 이하 <그림 7>에서 <그림 19>까지 제시하는 각 시기별, 기관별 핵심어의 네트워크도 시각화 자료의 식별을 위해 추출되는 단어가 80여 개가 되도록 연결횟수가 일정수준 이상이 되는 단어만 제시한다. 한편, 제시한 그림에서 핵심어의 연결중심성과 원의 크기는 비례한다.

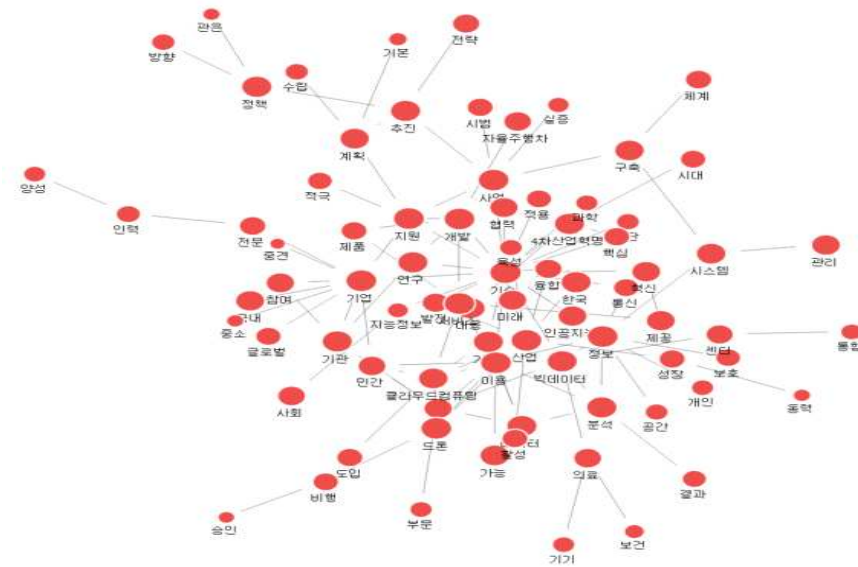
<표 2> 연결중심성 상위 30개 단어 (전체시기 : 2014. 1.~2018. 6.)

* 검색어로 사용한 단어는 밑줄 표시

| 순서 | 단어 | 연결중심성 | 순서 | 단어 | 연결중심성 |
|----|---------------|----------|----|----------------|----------|
| 1 | 기술 | 0.174651 | 16 | 기관 | 0.081358 |
| 2 | 이용 | 0.149437 | 17 | 정책 | 0.079845 |
| 3 | 정보 | 0.146579 | 18 | 한국 | 0.079173 |
| 4 | 산업 | 0.129434 | 19 | 기반 | 0.078501 |
| 5 | 개발 | 0.118003 | 20 | 추진 | 0.078164 |
| 6 | 드론 | 0.11649 | 21 | 나라 | 0.077828 |
| 7 | 서비스 | 0.111783 | 22 | <u>클라우드컴퓨팅</u> | 0.07682 |
| 8 | 기업 | 0.109934 | 23 | 안전 | 0.074466 |
| 9 | 사업 | 0.109262 | 24 | 정부 | 0.073458 |
| 10 | 지원 | 0.108758 | 25 | 계획 | 0.073122 |
| 11 | <u>빅데이터</u> | 0.105564 | 26 | 시스템 | 0.072617 |
| 12 | 분석 | 0.092957 | 27 | 대응 | 0.072281 |
| 13 | <u>4차산업혁명</u> | 0.090267 | 28 | 가능 | 0.072113 |
| 14 | 연구 | 0.086905 | 29 | 구축 | 0.070768 |
| 15 | 데이터 | 0.083207 | 30 | 관리 | 0.0706 |

<그림 6> 핵심어 간 네트워크 지도 (전체시기)

※ 명사 간 연결 횟수가 100회 이상인 경우 중 가장 큰 연결망



2014년 1월부터 2015년 12월까지의 시기는 이른바 ‘알파고 쇼크’ 이전 2년의 기간이다. 지능정보기술로 인한 사회변화에 대한 대중의 관심은 이세돌과 구글 딥마인드社 알파고의 바둑 대국이 있던 2016년 3월을 기점으로 상당히 증가하였다. 2014~2015년의 시기는 그 이전의 시기로서 아직까지 지능정보기술에 대한 정책적 논의가 활발하지 않다. 이에 분석대상 문서수가 이후시기에 비해 풍부하지 않다. 이 기간의 각 반기(半期)별 연결중심성 상위 30개 단어를 <표 3>과 같이 나타냈다. 각 반기별 핵심어 간 연결망을 시각화 결과는 <그림 7>에서 <그림 10>까지와 같다.

2014년 상반기에 연결중심성이 높은 단어는 ‘정보, 이용 서비스, 개발, 빅데이터’, 2014년 하반기에는 ‘정보, 이용, 기술, 클라우드 컴퓨팅, 빅데이터’, 2015년 상반기에는 ‘기술, 정보, 사물인터넷, 이용, 개발’, 2015년 하반기에는 ‘클라우드컴퓨팅, 정보, 빅데이터, 이용, 기술’ 등이다. 이 시기에는 ‘클라우드컴퓨팅, 빅데이터, 사물인터넷’ 등의 기술에 대한 관심이 높았다는 것을 알 수 있다.

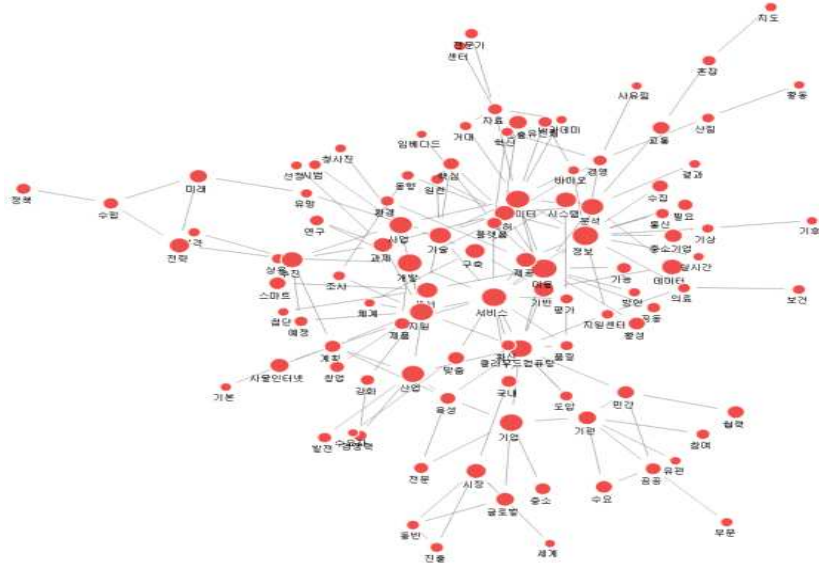
<표 3> 연결중심성 상위 30개 단어 (2014. 1.~2015. 12.)

* 검색어로 사용한 단어는 밑줄 표시

| | 2014년 상반기 (보도자료 26건) | | 2014년 하반기 (보도자료 28건) | | 2015년 상반기 (보도자료 52건) | | 2015년 하반기 (보도자료 71건) | |
|----|-------------------------|-------|-------------------------|-------|-------------------------|-------|-------------------------|-------|
| | 단어 | 중심성 | 단어 | 중심성 | 단어 | 중심성 | 단어 | 중심성 |
| 1 | 정보 | 0.145 | 정보 | 0.103 | 기술 | 0.114 | <u>클라우드</u> | 0.156 |
| 2 | 이용 | 0.120 | 이용 | 0.090 | 정보 | 0.103 | 정보 | 0.143 |
| 3 | 서비스 | 0.118 | 기술 | 0.083 | <u>사물인터넷</u> | 0.101 | <u>빅데이터</u> | 0.134 |
| 4 | 개발 | 0.098 | <u>클라우드</u> | 0.077 | 이용 | 0.097 | 이용 | 0.132 |
| 5 | <u>빅데이터</u> | 0.096 | <u>빅데이터</u> | 0.077 | 개발 | 0.092 | 기술 | 0.096 |
| 6 | <u>클라우드</u> | 0.092 | <u>사물인터넷</u> | 0.074 | 서비스 | 0.090 | 사업 | 0.086 |
| 7 | 지원 | 0.092 | 분석 | 0.067 | 사업 | 0.083 | 서비스 | 0.076 |
| 8 | 기업 | 0.092 | 서비스 | 0.062 | 산업 | 0.081 | 산업 | 0.073 |
| 9 | 산업 | 0.084 | 개발 | 0.048 | <u>클라우드</u> | 0.077 | 지원 | 0.067 |
| 10 | 사업 | 0.082 | 비만 | 0.041 | <u>빅데이터</u> | 0.071 | 정책 | 0.059 |
| 11 | 분석 | 0.078 | 보안 | 0.041 | 기업 | 0.071 | 분석 | 0.057 |
| 12 | 기술 | 0.074 | 산업 | 0.038 | 지원 | 0.070 | 구축 | 0.057 |
| 13 | 추진 | 0.062 | 사업 | 0.038 | 보안 | 0.057 | 기업 | 0.055 |
| 14 | 시스템 | 0.052 | 기업 | 0.036 | 협력 | 0.053 | <u>사물인터넷</u> | 0.054 |
| 15 | 센서 | 0.052 | 기반 | 0.036 | 계획 | 0.047 | 데이터 | 0.054 |
| 16 | 데이터 | 0.052 | 데이터 | 0.033 | 기반 | 0.046 | 계획 | 0.053 |
| 17 | 특허 | 0.050 | 안전 | 0.031 | 정책 | 0.045 | 정부 | 0.052 |
| 18 | 정부 | 0.048 | 초고 | 0.029 | 구축 | 0.045 | 추진 | 0.049 |
| 19 | 시장 | 0.048 | 교육 | 0.029 | <u>과학기술부</u> | 0.044 | 안전 | 0.049 |
| 20 | 기반 | 0.048 | 관리 | 0.029 | 추진 | 0.042 | 개발 | 0.049 |
| 21 | 구축 | 0.048 | 예측 | 0.028 | 실증 | 0.041 | 제공 | 0.045 |
| 22 | 제공 | 0.046 | 당뇨병 | 0.026 | 분석 | 0.040 | 기관 | 0.044 |
| 23 | 전략 | 0.046 | 지역 | 0.024 | 기관 | 0.038 | <u>드론</u> | 0.043 |
| 24 | 과제 | 0.044 | 임신 | 0.024 | 정부 | 0.035 | 기반 | 0.042 |
| 25 | <u>사물인터넷</u> | 0.042 | 스마트 | 0.024 | 제품 | 0.033 | 시장 | 0.040 |
| 26 | 창출 | 0.038 | 기관 | 0.024 | 제공 | 0.033 | 국내 | 0.040 |
| 27 | 기관 | 0.038 | 구축 | 0.024 | 핵심 | 0.032 | 보안 | 0.038 |
| 28 | 글로벌 | 0.038 | 통신 | 0.022 | 시스템 | 0.032 | 협력 | 0.037 |
| 29 | 중소기업 | 0.036 | 지원 | 0.022 | 과제 | 0.032 | 센터 | 0.037 |
| 30 | 미래 | 0.036 | 제품 | 0.022 | 가능 | 0.032 | 필요 | 0.035 |

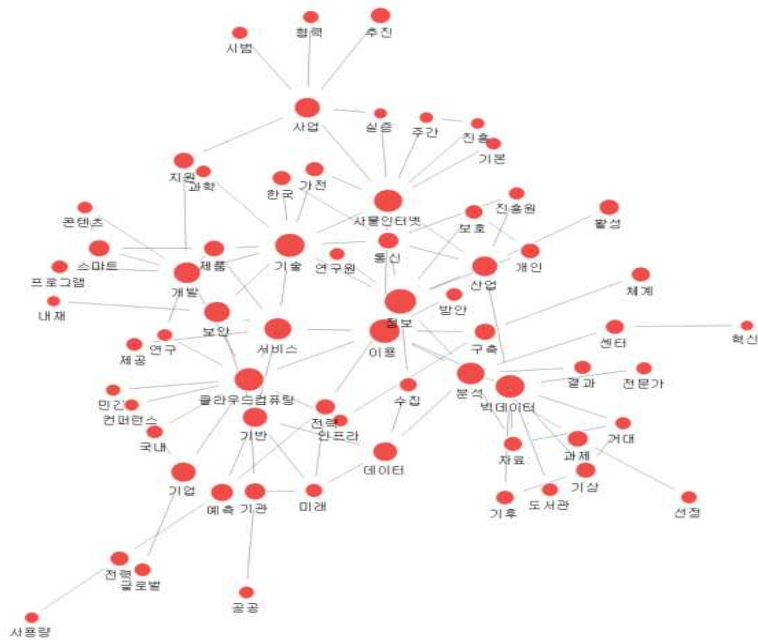
<그림 7> 핵심어 간 네트워크 지도 ('14년 상반기)

※ 명사 간 연결 횟수가 5회 이상인 경우 중 가장 큰 연결망



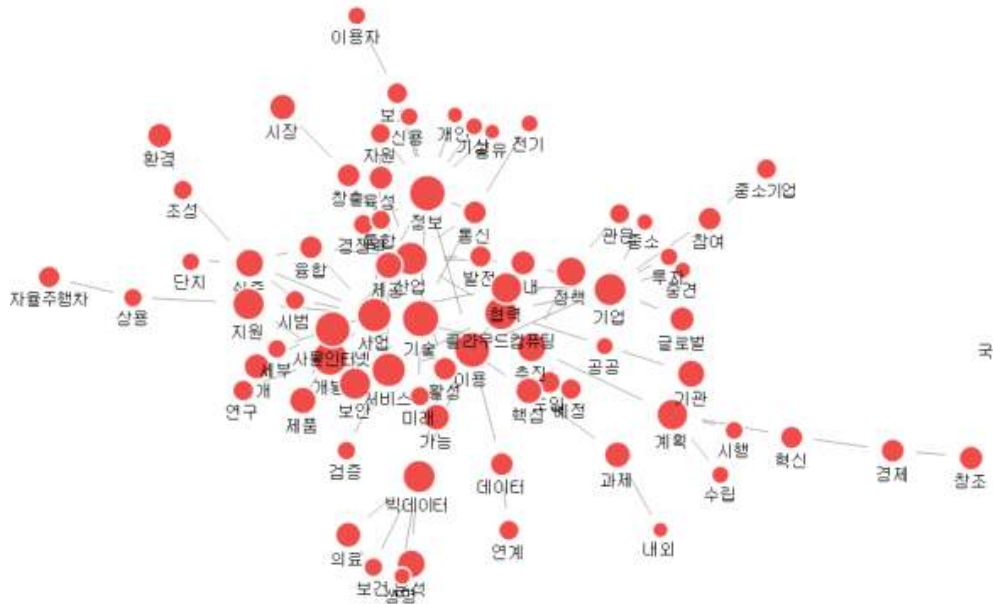
<그림 8> 핵심어 간 네트워크 지도 ('14년 하반기)

※ 명사 간 연결 횟수가 5회 이상인 경우 중 가장 큰 연결망



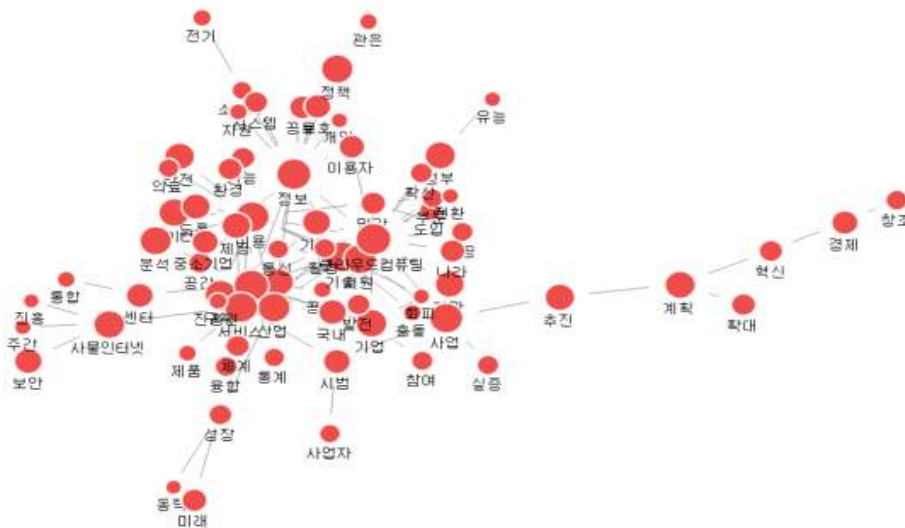
<그림 9> 핵심어 간 네트워크 지도 ('15년 상반기)

※ 명사 간 연결 횟수가 7회 이상인 경우 중 가장 큰 연결망



〈그림 10〉 핵심어 간 네트워크 지도 ('15년 하반기)

※ 명사 간 연결 횟수가 10회 이상인 경우 중 가장 큰 연결망



2016년 1월부터 2016년 12월까지의 시기는 이세돌과 구글 딥마인드社 알파고의 바둑 대국이후, 지능정보기술에 대한 사회적인 관심이 증가하는 시기이다. 대중에게 상당한 충격을 가져다준 이벤트였다. 당시 미래창조과학부는 2016년 3월에 ‘지능정보산업 발전전략’을 발표하였으며, 2016년 4월에 ‘지능정보사회 중장기 종합대책 추진 계획’이 국무회의에 상정되었다. 이에 따라 2016년 5월 범부처 협력체계인 민관합동추진협의회가 설립되고, 2016년 9월 지능정보사회 추진단이 설립되었다. 이후, ‘지능정보사회 추진 민관 합동 컨퍼런스’를 거쳐 2016년 12월 범정부 종합대책인 ‘지능정보사회 중장기 종합대책’이 마련되었다. 이 기간의 각 반기(半期)별 연결중심성 상위 30개 단어는 <표 4>와 같다. 한편, 각 반기별 핵심어 간 연결망을 시각화 결과는 <그림 11>에서 <그림 12>까지와 같다.

2016년 상반기에 연결중심성이 높은 단어는 ‘기술, 이용, 정보, 지원, 산업’, 2016년 하반기에는 ‘기술, 이용, 산업, 정보, 개발’ 등이다. 이 시기에는 이전 시기에 관심이 높았던 ‘클라우드컴퓨팅, 빅데이터, 사물인터넷’ 등 외에 ‘드론, 자율주행차, 스마트시티’ 등에 대한 관심이 상당히 높아졌음을 알 수 있다. ‘클라우드컴퓨팅, 빅데이터, 사물인터넷’ 등의 기술은 다른 기술에 응용·활용될 수 있는 기저(基底) 기술이며, ‘드론, 자율주행차, 스마트시티’는 위 기술을 활용하여 구체적으로 활용하는 형태를 보여주는 기술이다. 이와 같은 측면에서 이전에 추상적으로 논의된 지능정보기술에 대한 관심이 산업적으로 구체화된다는 점에서 의미가 있다고 할 수 있다.

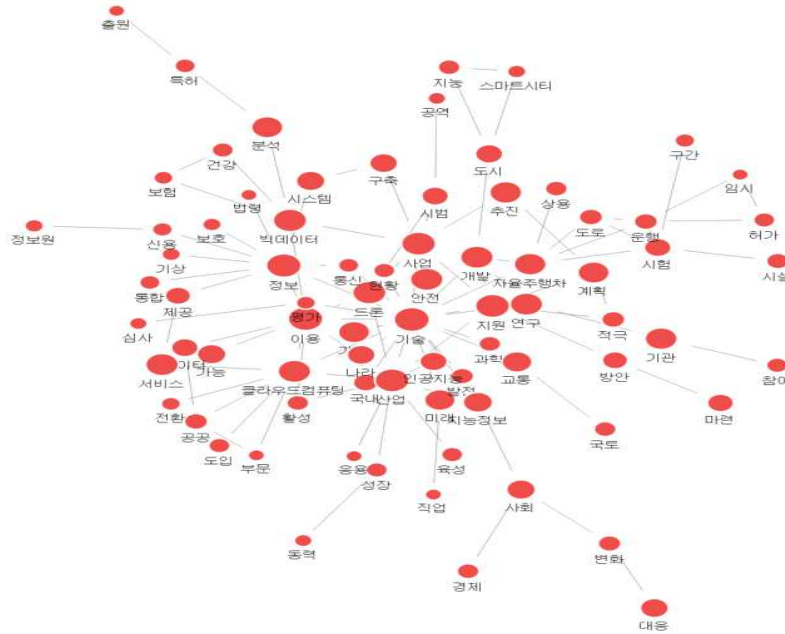
<표 4> 연결중심성 상위 30개 단어 (2016. 1.~2016. 12.)

* 검색어로 사용한 단어는 밑줄 표시

| | 2016년 상반기 (보도자료 101건) | | 2016년 하반기 (보도자료 158건) | |
|----|--------------------------|-------|--------------------------|-------|
| | 단어 | 중심성 | 단어 | 중심성 |
| 1 | 기술 | 0.150 | 기술 | 0.137 |
| 2 | 이용 | 0.097 | 이용 | 0.122 |
| 3 | 정보 | 0.096 | 산업 | 0.107 |
| 4 | 지원 | 0.088 | 정보 | 0.103 |
| 5 | 산업 | 0.087 | 개발 | 0.092 |
| 6 | 사업 | 0.085 | <u>드론</u> | 0.086 |
| 7 | <u>빅데이터</u> | 0.074 | <u>빅데이터</u> | 0.083 |
| 8 | <u>드론</u> | 0.074 | 서비스 | 0.078 |
| 9 | <u>클라우드컴퓨팅</u> | 0.064 | 지원 | 0.070 |
| 10 | 개발 | 0.061 | 사업 | 0.067 |
| 11 | 연구 | 0.053 | 기업 | 0.067 |
| 12 | 안전 | 0.053 | 연구 | 0.058 |
| 13 | 서비스 | 0.053 | 기관 | 0.057 |
| 14 | <u>자율주행차</u> | 0.049 | 분석 | 0.056 |
| 15 | 추진 | 0.047 | 추진 | 0.055 |
| 16 | 기관 | 0.047 | 공공 | 0.053 |
| 17 | 분석 | 0.047 | 기반 | 0.053 |
| 18 | 기업 | 0.046 | 정책 | 0.052 |
| 19 | 계획 | 0.044 | 계획 | 0.052 |
| 20 | 정책 | 0.043 | 안전 | 0.051 |
| 21 | 미래 | 0.042 | 구축 | 0.048 |
| 22 | 교통 | 0.040 | 시스템 | 0.047 |
| 23 | 관리 | 0.040 | 데이터 | 0.047 |
| 24 | 자동차 | 0.039 | 도시 | 0.046 |
| 25 | 국민 | 0.038 | <u>스마트시티</u> | 0.045 |
| 26 | <u>지능정보</u> | 0.038 | 미래 | 0.043 |
| 27 | 가능 | 0.037 | <u>클라우드컴퓨팅</u> | 0.042 |
| 28 | 정부 | 0.036 | 정부 | 0.041 |
| 29 | 사회 | 0.036 | 가능 | 0.041 |
| 30 | 시스템 | 0.036 | 훈련 | 0.040 |

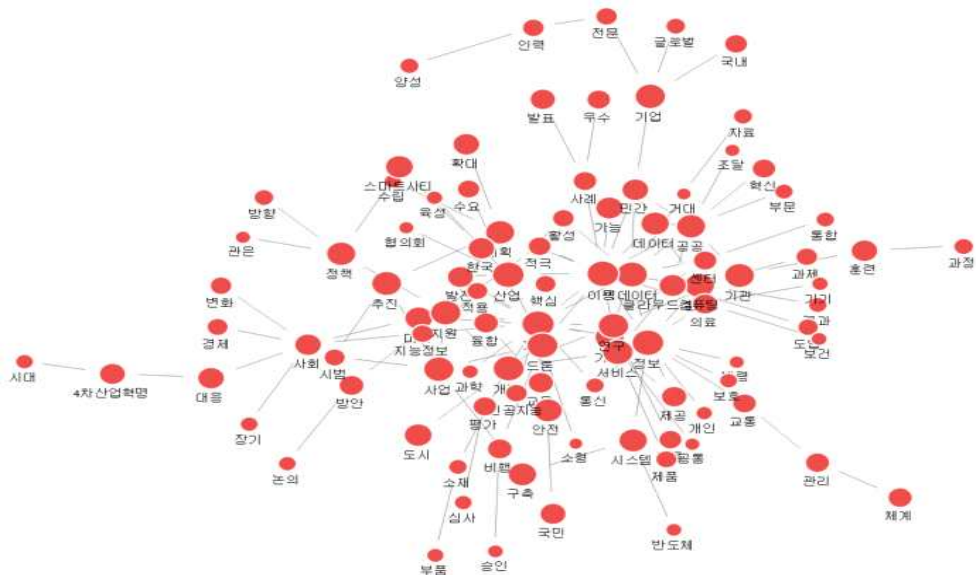
〈그림 11〉 핵심어 간 네트워크 지도 ('16년 상반기)

※ 명사 간 연결 횟수가 12회 이상인 경우 중 가장 큰 연결망



〈그림 12〉 핵심어 간 네트워크 지도 ('16년 하반기)

※ 명사 간 연결 횟수가 15회 이상인 경우 중 가장 큰 연결망



2017년 1월부터 2018년 6월까지의 시기의 각 반기(半期)별 연결 중심성 상위 30개 단어는 <표 5>에 제시하였다. 2016년 12월 범정부 종합대책인 ‘지능정보사회 중장기 종합대책’이 마련된 이후, 관련 정책적 논의가 활발해지면서 보도자료의 수도 이전 시기에 비해 상당히 증가하였다. 지능정보기술에 대한 사회적 관심이 높아지면서 2017년 5월에는 대선(大選)에서 주요 후보들의 공약으로도 나타났다. 새 정부 출범 이후, 2017년 8월 대통령 직속 4차 산업혁명위원회가 출범하고, 2017년 12월 ‘사람중심의 4차 산업혁명 대응계획’ 발표 등의 정책적 논의가 있었다. 각 반기별 핵심어 간 연결망을 시각화 결과는 <그림 7>에서 <그림 15>까지와 같다.

2017년 상반기에 연결중심성이 높은 단어는 ‘기술, 산업, 기업, 지원, 이용’, 2017년 하반기에는 ‘기술, 정보, 산업, 4차산업혁명, 드론’, 2018년 상반기에는 ‘기술, 이용, 드론, 사업, 개발’ 등이다.

이 시기의 이전 시기와 구분되는 특징으로는 ‘4차 산업혁명’에 대한 언급 증가를 꼽을 수 있다. 학계의 4차 산업혁명 용어 사용에 대한 회의적인 학계 시각(김동욱, 2018; 성욱준·황성수, 2017)에도 불구하고, 지능정보정책을 설명하는 대중적인 키워드가 되었다. ‘4차 산업혁명’이라는 용어는 2016년 다보스 포럼에서 K. Schwab이 사용한 이래, 2016년 12월 ‘제4차 산업 혁명에 대응한 지능정보사회 중장기 종합대책’ (관계부처 합동, 2016)이 발표되면서 정부 문서에도 공식적으로 사용되었고, 2017년 8월 대통령 직속기구로 ‘4차 산업혁명위원회’가 출범하면서 위원회 명칭에 까지 등장하였다. 또, 이 시기에 비로소 ‘인공지능’이 연결중심성 상위 30개 단어로 등장하였다.

한편, 이 시기에는 이전까지 계속하여 증가하던 관련 보도자료의 문건 숫자가 2017년 하반기, 2018년 상반기로 가면서 점차 줄어드는 양상도 확인할 수 있다.

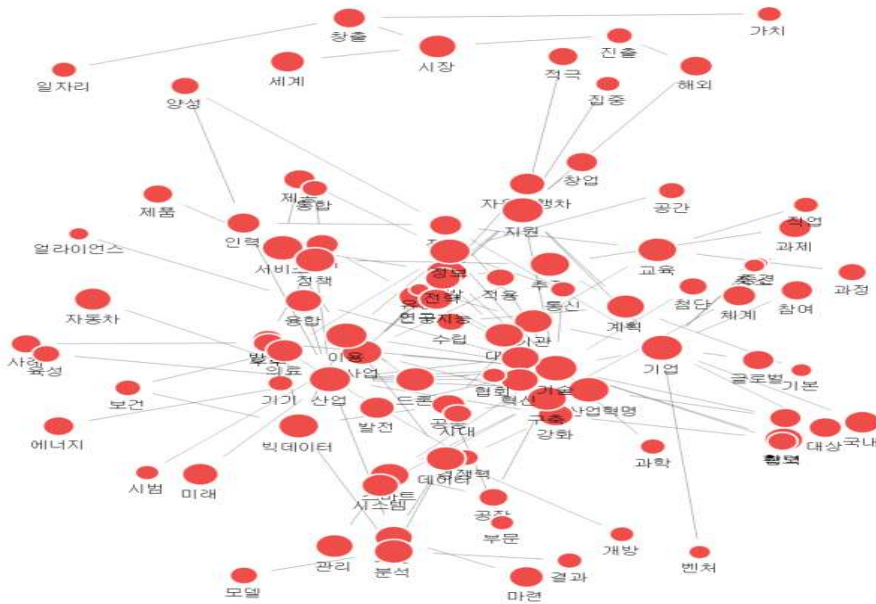
<표 5> 연결중심성 상위 30개 단어 (2017. 1.~2018. 6.)

* 검색어로 사용한 단어는 밑줄 표시

| | 2017년 상반기 (보도자료 258건) | | 2017년 하반기 (보도자료 241건) | | 2018년 상반기 (보도자료 185건) | |
|----|--------------------------|-------|--------------------------|-------|--------------------------|-------|
| | 단어 | 중심성 | 단어 | 중심성 | 단어 | 중심성 |
| 1 | 기술 | 0.156 | 기술 | 0.169 | 기술 | 0.165 |
| 2 | 산업 | 0.139 | 정보 | 0.114 | 이용 | 0.130 |
| 3 | 기업 | 0.108 | 산업 | 0.108 | <u>드론</u> | 0.115 |
| 4 | 지원 | 0.100 | <u>4차산업혁명</u> | 0.104 | 사업 | 0.102 |
| 5 | 이용 | 0.098 | <u>드론</u> | 0.095 | 개발 | 0.102 |
| 6 | 개발 | 0.095 | 이용 | 0.093 | 산업 | 0.094 |
| 7 | <u>4차산업혁명</u> | 0.092 | 기업 | 0.092 | 지원 | 0.092 |
| 8 | 서비스 | 0.090 | 지원 | 0.090 | 정보 | 0.090 |
| 9 | 정보 | 0.082 | 서비스 | 0.082 | 서비스 | 0.081 |
| 10 | 사업 | 0.073 | 개발 | 0.079 | 기관 | 0.078 |
| 11 | <u>빅데이터</u> | 0.071 | 연구 | 0.073 | 기업 | 0.077 |
| 12 | 정책 | 0.070 | 정책 | 0.067 | 추진 | 0.074 |
| 13 | 추진 | 0.066 | 사업 | 0.067 | 혁신 | 0.072 |
| 14 | 스마트 | 0.062 | 한국 | 0.065 | 연구 | 0.072 |
| 15 | 대응 | 0.061 | 기반 | 0.065 | 분석 | 0.071 |
| 16 | 교육 | 0.061 | <u>빅데이터</u> | 0.062 | 계획 | 0.069 |
| 17 | <u>드론</u> | 0.058 | <u>스마트시티</u> | 0.061 | <u>빅데이터</u> | 0.068 |
| 18 | 분석 | 0.056 | 추진 | 0.058 | 안전 | 0.066 |
| 19 | 데이터 | 0.056 | 혁신 | 0.056 | <u>4차산업혁명</u> | 0.066 |
| 20 | 구축 | 0.056 | 계획 | 0.055 | <u>인공지능</u> | 0.063 |
| 21 | 기관 | 0.055 | 교육 | 0.052 | 구축 | 0.060 |
| 22 | 계획 | 0.055 | 환경 | 0.051 | 정책 | 0.059 |
| 23 | 협력 | 0.054 | 대응 | 0.051 | <u>스마트시티</u> | 0.057 |
| 24 | 한국 | 0.054 | 안전 | 0.051 | 나라 | 0.056 |
| 25 | 정부 | 0.054 | 나라 | 0.051 | 기반 | 0.056 |
| 26 | 나라 | 0.054 | 데이터 | 0.050 | 한국 | 0.054 |
| 27 | 연구 | 0.053 | 분석 | 0.049 | 도시 | 0.053 |
| 28 | 확대 | 0.050 | 구축 | 0.048 | 데이터 | 0.053 |
| 29 | 안전 | 0.049 | 기관 | 0.048 | 공공 | 0.051 |
| 30 | 혁신 | 0.048 | 관리 | 0.046 | 정부 | 0.050 |

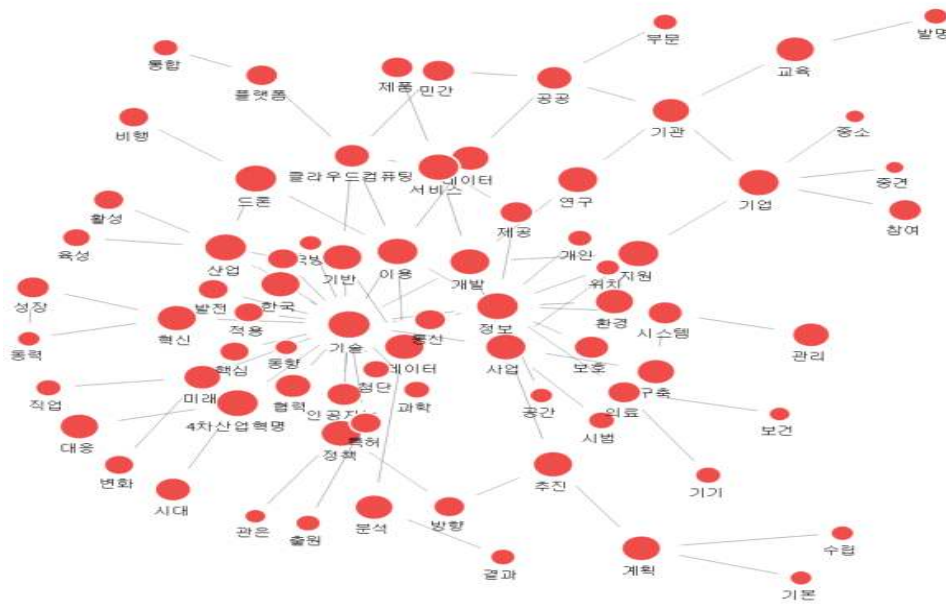
<그림 13> 핵심어 간 네트워크 지도 ('17년 상반기)

※ 명사 간 연결 횟수가 25회 이상인 경우 중 가장 큰 연결망



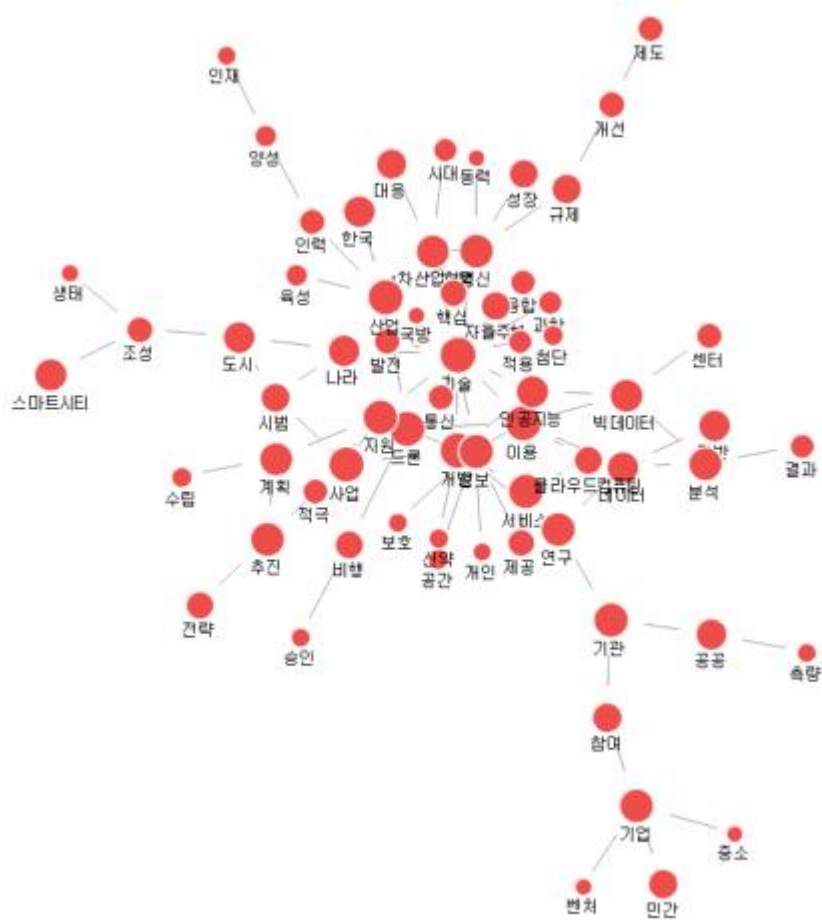
<그림 14> 핵심어 간 네트워크 지도 ('17년 하반기)

※ 명사 간 연결 횟수가 25회 이상인 경우 중 가장 큰 연결망



<그림 15> 핵심어 간 네트워크 지도 ('18년 상반기)

※ 명사 간 연결 횟수가 25회 이상인 경우 중 가장 큰 연결망



이상 시기별 핵심어 분석을 통해 확인할 수 있는 사항은 다음과 같이 정리할 수 있다.

먼저, 각 시기별 지능정보기술 활용의 초점이 달라졌다. 분석 기간 중 초기에 해당하는 시기(2014년~2015년)의 경우, 지능정보기술 활용의 초점이 ‘업무와 서비스를 지원’ 하는 측면에 있었으나, 시간이 지나면서 ‘산업과 기업에 활용’ 하는 측면이 보다 강조되었다.

둘째, 각 시기별 자주 언급되는 지능정보기술이 다르다. 2016년 이후에는 이전에 연결중심성 상위 단어로 등장하지 않던 ‘드론, 스마트시티, 4차 산업혁명, 인공지능’ 이 등장하였고, 분석 기간 중 초기에 해당하는 시기(2014년~2015년)에 연결중심성 상위에 속하던 ‘클라우드 컴퓨팅, 빅데이터, 사물인터넷’ 의 비중이 상대적으로 줄어들었다는 점을 알 수 있다.

셋째, ‘4차 산업혁명’ 용어의 등장이다. 2014년~2015년에 없었던 ‘4차 산업혁명’ 이라는 용어는 2016년 이후 활발하게 사용되기 시작하였음을 확인할 수 있다.

2. 주요 기관별 비교

지능정보정책에 관한 키워드를 통해 수집된 보도자료는 과학기술정보통신부(271건), 국토교통부(254건), 산업통상자원부(122건), 행정안전부(112건) 순으로 많았다. 특히, 위 네 개의 중앙행정기관의 보도자료가 전체 보도자료의 66.4%를 차지하고 있어 지능정보정책에 관한 보도자료가 특정 부처에 집중되어 배포되고 있음을 알 수 있다. 이에 보도자료가 가장 많이 수집된 상위 네 개 중앙행정기관에 대해 분석을 진행하였다.

주요 행정기관별 보도자료에서 수집된 단어 중 연결중심성이 높은 단어 상위 30개를 살펴보면 <표 6>과 같다.

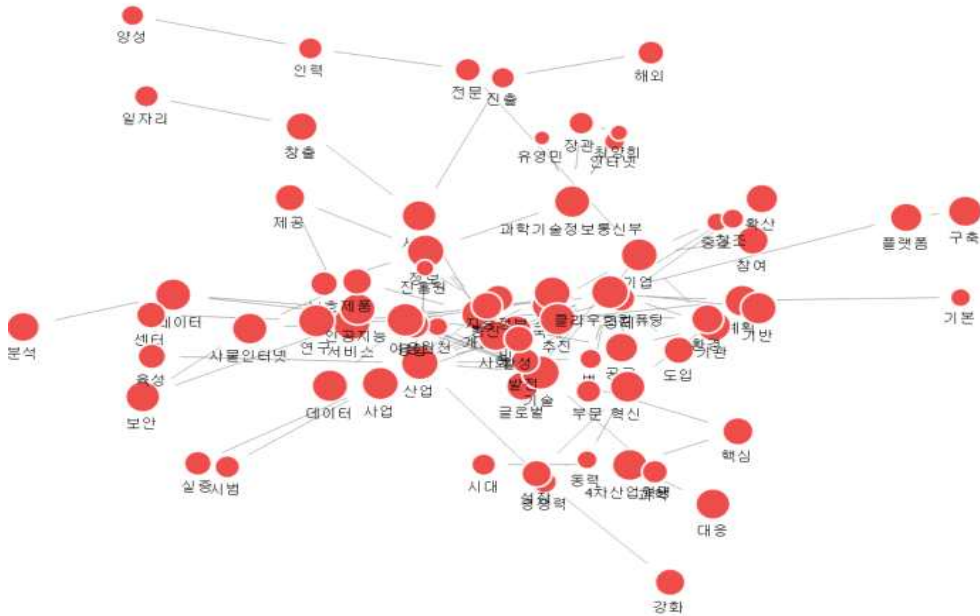
<표 6> 연결중심성 상위 30개 단어 (주요기관)

* 검색어로 사용한 단어는 밑줄 표시

| | 과학기술정보통신부 (보도자료 271건) | | 산업통상자원부 (보도자료 122건) | | 국토교통부 (보도자료 254건) | | 행정안전부 (보도자료 112건) | |
|----|--------------------------|-------|------------------------|-------|----------------------|-------|----------------------|-------|
| | 단어 | 중심성 | 단어 | 중심성 | 단어 | 중심성 | 단어 | 중심성 |
| 1 | 기술 | 0.175 | 산업 | 0.188 | <u>드론</u> | 0.160 | 분석 | 0.170 |
| 2 | <u>클라우드</u> | 0.144 | 기술 | 0.176 | 기술 | 0.114 | <u>빅데이터</u> | 0.166 |
| 3 | 서비스 | 0.136 | 기업 | 0.146 | 이용 | 0.112 | 이용 | 0.159 |
| 4 | 이용 | 0.131 | 지원 | 0.108 | 사업 | 0.101 | 정보 | 0.127 |
| 5 | 정보 | 0.125 | 개발 | 0.098 | 도시 | 0.089 | 데이터 | 0.112 |
| 6 | 산업 | 0.117 | <u>4차산업혁명</u> | 0.080 | 안전 | 0.085 | 정부 | 0.110 |
| 7 | 지원 | 0.111 | 사업 | 0.078 | 지원 | 0.080 | 서비스 | 0.101 |
| 8 | 개발 | 0.111 | 스마트 | 0.074 | 개발 | 0.078 | 공공 | 0.097 |
| 9 | 기업 | 0.104 | <u>드론</u> | 0.071 | <u>스마트시티</u> | 0.076 | <u>클라우드</u> | 0.090 |
| 10 | 연구 | 0.092 | 서비스 | 0.068 | 추진 | 0.070 | 기관 | 0.090 |
| 11 | 사업 | 0.091 | 협력 | 0.066 | 계획 | 0.067 | 국민 | 0.082 |
| 12 | <u>과학기술부</u> | 0.090 | 혁신 | 0.066 | 구축 | 0.065 | 기술 | 0.073 |
| 13 | <u>빅데이터</u> | 0.090 | 이용 | 0.066 | <u>국토교통부</u> | 0.064 | 지원 | 0.069 |
| 14 | 데이터 | 0.084 | 융합 | 0.063 | <u>자율주행</u> | 0.062 | 추진 | 0.067 |
| 15 | 추진 | 0.080 | 추진 | 0.057 | 산업 | 0.061 | 기반 | 0.066 |
| 16 | 기반 | 0.080 | 산업통상부 | 0.056 | 비행 | 0.061 | 사업 | 0.065 |
| 17 | <u>사물인터넷</u> | 0.080 | 한국 | 0.052 | 정보 | 0.059 | 센터 | 0.062 |
| 18 | 혁신 | 0.077 | 로봇 | 0.052 | 혁신 | 0.057 | <u>행정안전부</u> | 0.058 |
| 19 | <u>인공지능</u> | 0.069 | 시장 | 0.052 | 시범 | 0.051 | 부처 | 0.058 |
| 20 | <u>4차산업혁명</u> | 0.067 | 시스템 | 0.052 | 기업 | 0.051 | 행정 | 0.057 |
| 21 | 계획 | 0.067 | 정책 | 0.051 | 지역 | 0.050 | 정책 | 0.056 |
| 22 | 시장 | 0.063 | 정보 | 0.050 | 관리 | 0.049 | 업무 | 0.056 |
| 23 | 보안 | 0.062 | 계획 | 0.049 | 서비스 | 0.046 | 안전 | 0.056 |
| 24 | 대응 | 0.060 | 국내 | 0.048 | 항공 | 0.045 | 계획 | 0.056 |
| 25 | 정책 | 0.059 | 에너지 | 0.048 | 시험 | 0.045 | 시스템 | 0.055 |
| 26 | 구축 | 0.058 | 기반 | 0.048 | 시스템 | 0.045 | 구축 | 0.054 |
| 27 | 기관 | 0.058 | 표준 | 0.047 | <u>자율주행차</u> | 0.044 | 과제 | 0.049 |
| 28 | <u>스마트시티</u> | 0.057 | 대응 | 0.046 | 인프라 | 0.044 | 지자체 | 0.045 |
| 29 | 사회 | 0.057 | 확대 | 0.045 | 기반 | 0.044 | 재난 | 0.045 |
| 30 | 국내 | 0.055 | 연구 | 0.045 | 민간 | 0.043 | 운영 | 0.045 |

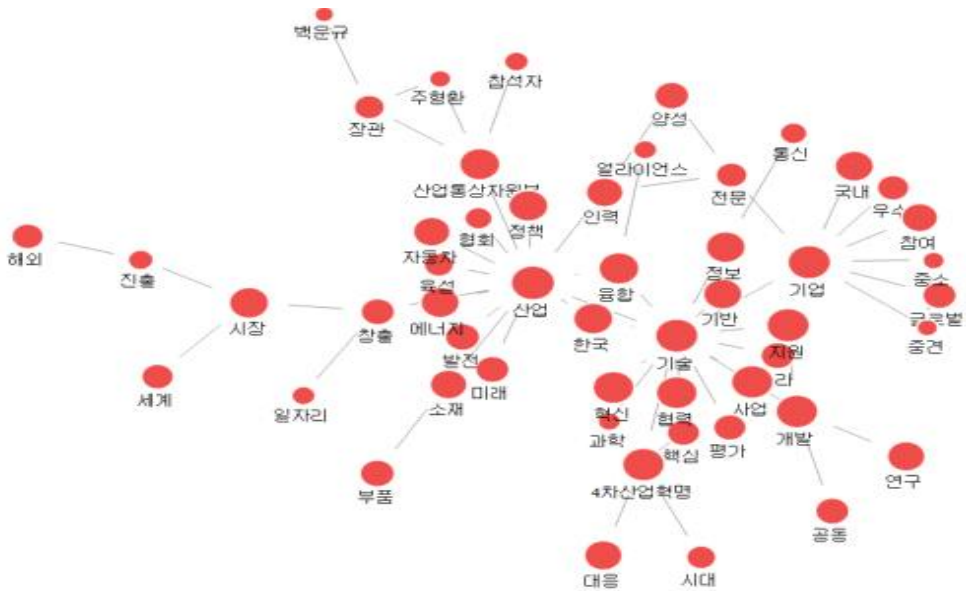
<그림 16> 핵심어 간 네트워크 지도 (과학기술정보통신부)

※ 명사 간 연결 횟수가 33회 이상인 경우 중 가장 큰 연결망



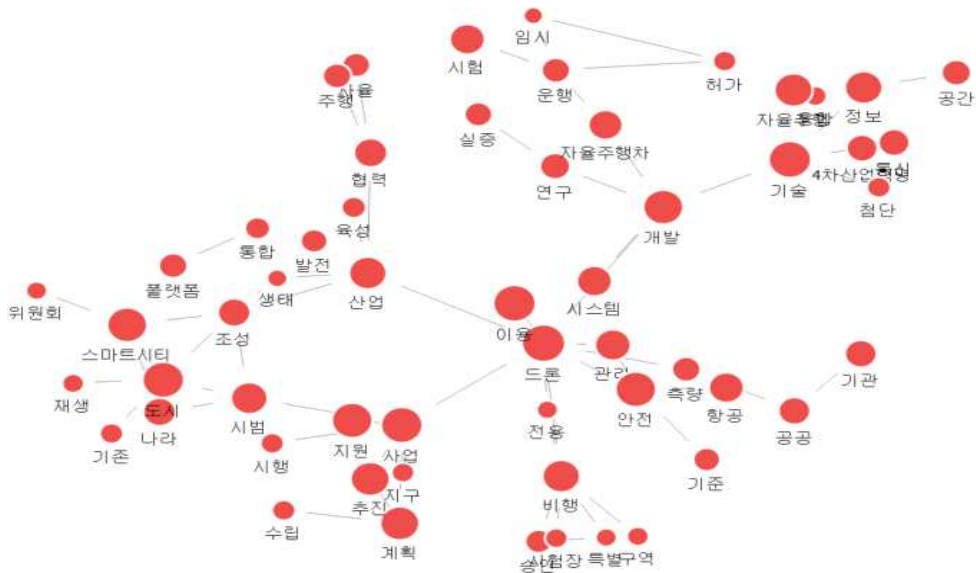
<그림 17> 핵심어 간 네트워크 지도 (산업통상자원부)

※ 명사 간 연결 횟수가 33회 이상인 경우 중 가장 큰 연결망



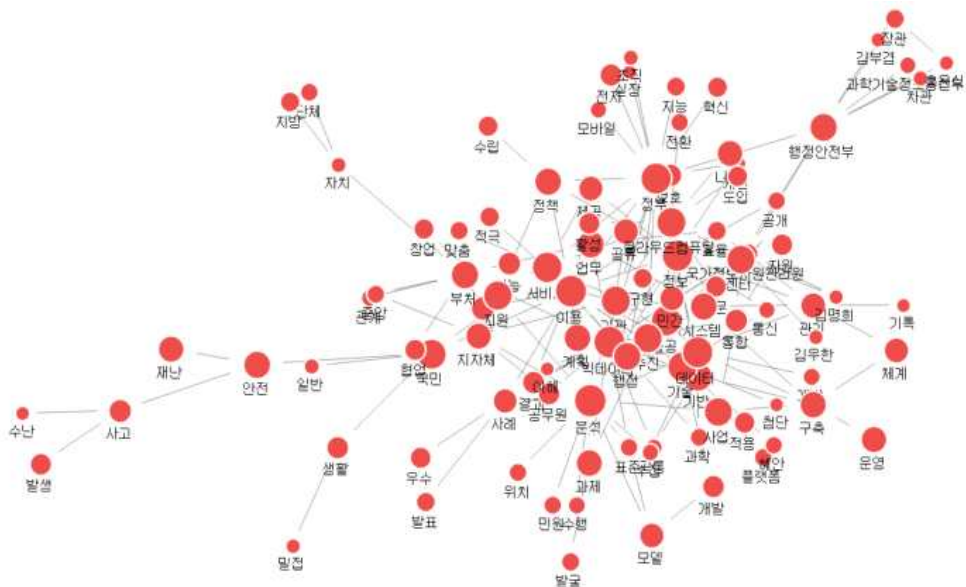
〈그림 18〉 핵심어 간 네트워크 지도 (국토교통부)

※ 명사 간 연결 횟수가 13회 이상인 경우 중 가장 큰 연결망



〈그림 19〉 핵심어 간 네트워크 지도 (행정안전부)

※ 명사 간 연결 횟수가 13회 이상인 경우 중 가장 큰 연결망



각 부처별로 살펴보면 다음과 같이 정리할 수 있다. 먼저, 과학기술정보통신부는 ‘기술, 클라우드 컴퓨팅, 서비스, 이용, 정보’ 등을 주요 핵심어로 사용하였다. 전체 행정기관을 통해 분석한 자료와 비교할 때, 상대적으로 ‘클라우드 컴퓨팅, 빅데이터, 사물인터넷, 혁신, 국내’ 등이 강조된다. 산업통상자원부는 ‘산업, 기술, 기업, 지원, 개발’ 등을 주요 핵심어로 사용하였다. 상대적으로 ‘기업, 스마트, 4차산업혁명, 융합, 혁신’ 등이 강조된다. 국토교통부는 ‘드론, 기술, 이용, 사업, 도시’ 등을 주요 핵심어로 사용하였다. 상대적으로 ‘드론, 사업, 스마트시티, 안전, 도시’ 등이 강조된다. 마지막으로 행정안전부는 ‘분석, 빅데이터, 이용, 정보, 데이터’ 등을 주요 핵심어로 사용하였다. 상대적으로 ‘빅데이터, 분석, 데이터, 공공, 클라우드 컴퓨팅’ 등이 강조된다.

주요 부처별 빈도 분석을 통해 행정안전부와 나머지 세 개(과학기술정보통신부, 국토교통부, 산업통상자원부) 기관의 추출된 단어 특성이 구분됨을 알 수 있다. 구체적으로 행정안전부를 제외한 나머지 세 개 기관이 상대적으로 ‘산업, 기업, 개발’ 등의 단어의 연결중심성이 높게 나타나고 있음을 알 수 있다. 위 기관들이 지능정보정책을 주로 산업적인 활용에 연관 짓고 있다고 해석할 수 있다. 다만, 과학기술정보통신부는 ‘클라우드 컴퓨팅’을, 국토교통부는 ‘드론, 스마트시티’를 주요 핵심어로 사용하는 등, 중점을 두고 있는 지능정보기술에는 기관의 차이가 있다. 한편, 산업통상자원부는 특정 기술보다도 ‘4차 산업혁명, 스마트, 융합, 혁신’ 등 단어의 연결중심성이 높게 나타나고 있음을 알 수 있다.

반면, 행정안전부는 ‘공공, 정부, 기관’ 등의 연결중심성이 높게 나타나고 있으며, 핵심어로 사용되는 지능정보기술은 ‘빅데이터, 클라우드 컴퓨팅’ 등과 같이 업무에 활용·지원될 수 있는 기술들이다.

이를 통해 볼 때, 행정안전부는 정부나 공공기관에서 빅데이터를 분석하고, 클라우드를 활용하여, 업무시스템을 개선하는 측면에 초점을 두고 있음을 알 수 있다. 이는 행정안전부가 지원·통제 기능을 수행하는 참모(staff) 부처의 성격을 지니고, 과학기술정보통신부, 국토교통부, 산업통상자원부가 산업 진흥 등의 역할을 수행하는 경제 부처의 성격을 지닌다는 것이 핵심어 분석에서도 나타났다고 볼 수 있다.

제 3 절 단어분포의 집중도 분석

1. 시기별 비교

인구분포와 소득분포와의 관계를 나타내는 수치로서 소득분배의 불평등도를 나타내는 지니계수를 단어의 종류와 단어와 연결된 링크의 숫자의 관계에 적용하여 관계를 나타내었다. 이렇게 도출한 수치를 본 연구에서는 ‘단어분포의 집중도’로 해석하였다. 핵심어 분석과 마찬가지로 수집 자료의 특성상 무의미하게 반복되는 단어를 제외하고, 의미 있는 분석을 위해 명사를 노드로 하는 의미연결망(window size=3)을 도출하고, 단어가 2회 이상 함께 등장한 경우를 대상으로 하였다. 또, 빈도를 기반으로 하지 않고, 연결된 링크 숫자를 중심으로 도출하였다. 결과는 <표 7>과 같다.

분석 대상 기간을 6개월로 나누어 살펴보면, 시간이 지날수록 단어분포의 집중도가 커지는 경향이 있음을 알 수 있다. 즉, 시간이 지날수록 연결중심성이 높은 핵심어가 보도자료에서 차지하는 비중이 더 커지고 있음을 확인할 수 있다.

한편, 이는 각 시기 내의 단어 분포에서 상대적인 분포를 나타내는 것이기 때문에 단어분포의 집중도가 더 높은 2016년 이후 시기가 그 이전 시기에 비해 더 적은 주제를 담는다고 해석 할 수는 없다. 특히, 뒤의 시기로 갈수록 수집된 문서와 단어 숫자가 많아지는 경향이 나타나기 때문에 시간이 갈수록 관련 논의 자체는 더 많아지지만, 특정 분야의 집중되는 정도는 더 커진다고 해석 할 수도 있다.

이와 같이 도출된 수치를 비판적으로 해석하면 시간일 지날수록 정책적 관심이 특정 주제에 한정된다고 볼 수도 있다. 반면, 긍정적으로 해석하면 시간이 지날수록 특정 주제에 관한 정책적 관심이 구체화되어 집중된다고 볼 수 있다. 당초 예상과도 부합하는 결과라고 할 수

있다.

<표 7> 지니계수를 활용한 단어분포의 집중도(시기별)

| | 전체 | '14년 상반기 | '14년 하반기 | '15년 상반기 | '15년 하반기 |
|--------------|-------------|-------------|-------------|-------------|-------------|
| 단어분포의 집중도 | 0.789 | 0.585 | 0.531 | 0.592 | 0.618 |
| | '16년 상반기 | '16년 하반기 | '17년 상반기 | '17년 상반기 | '18년 하반기 |
| 단어분포의 집중도 | 0.64 | 0.685 | 0.73 | 0.706 | 0.706 |

2. 기관별 비교

시기별 비교에서도 살펴본 바와 같이 인구분포와 소득분포와의 관계를 나타내는 수치로서 소득분배의 불평등도를 나타내는 지니계수를 활용하여 단어의 종류와 단어에 연결된 링크의 숫자의 관계에 적용하여 단어분포의 집중도 관계를 나타내어 본다. 그 결과 단어 집중도는 과학기술정보통신부(0.726), 산업통상자원부(0.706), 행정안전부(0.671), 국토교통부(0.635) 순으로 나타났다. 즉, 단어분포의 집중도를 통해 살펴 봤을 때, 과학기술정보통신부, 산업통상자원부, 행정안전부, 국토교통부 순으로 특정 단어가 집중되어있는 정도가 크다고 할 수 있다.

한편, 시간의 흐름에 따른 변화를 살펴보아도 각 개별 기관별로도 정도의 차이는 있지만, 전체 기관을 대상으로 한 결과와 비슷하게 시간이 지날수록 단어분포의 집중도가 커지고 있음을 확인할 수 있다.⁹⁾

9) 단, 2014년의 경우, 개별 부처 단위 표본이 적어서 전체적인 수치 경향과 매우 동떨어져있음을 확인할 수 있다.

본 분석결과는 이는 기술과 산업 일반을 다루는 과학기술정보통신부와 산업통상자원부의 단어집중도가 낮을 것이라는 당초 예상과 다른 결과이다. 지능정보정책 전반에 대한 주무기관인 과학기술정보통신부와 관련 산업 정책에 대한 주무기관인 산업통상자원부의 단어 분포가 보다 집중적인 것은 지능정보정책에 관한 보도자료가 정책을 다루는 주제가 명확하기 때문으로 생각할 수 있다. 반면, 특정 주제에만 집중되어 다양한 주제를 다루지 못한다고 비판할 수도 있다.

상대적으로 단어분포가 고른 편인 행정안전부는 각 부처의 지원·통제 기능을 수행하므로 사용 단어가 더 다양하다고 해석할 수 있다. 국토교통부는 ‘드론, 스마트시티, 자율주행차’ 등의 건설, 교통과 관련된 특정 산업을 다루고 있음에도 단어 분포가 상대적으로 고른 편으로 나타났다. 건설, 교통 분야 지능정보정책과 관련한 세부주제가 다양하다고 해석할 수 있다.

<표 8> 지니계수를 활용한 단어분포의 집중도(기관별)

| | 전체 | '14년 상반기 | '14년 하반기* | '15년 상반기 | '15년 하반기 |
|-------|-------------|-------------|--------------|-------------|-------------|
| 전체 | 0.789 | 0.585 | 0.531 | 0.592 | 0.618 |
| 과기정통부 | 0.726 | 0.512 | 0.505 | 0.573 | 0.572 |
| 산업통상부 | 0.706 | 0.385 | (자료없음) | 0.355 | 0.456 |
| 국토교통부 | 0.635 | 0.789 | (자료없음) | 0.431 | 0.506 |
| 행정안전부 | 0.671 | 0.385 | 0.404 | 0.42 | 0.531 |
| | '16년 상반기 | '16년 하반기 | '17년 상반기 | '17년 상반기 | '18년 하반기 |
| 전체 | 0.64 | 0.685 | 0.73 | 0.706 | 0.706 |
| 과기정통부 | 0.554 | 0.548 | 0.553 | 0.651 | 0.62 |
| 산업통상부 | 0.44 | 0.555 | 0.664 | 0.555 | 0.576 |
| 국토교통부 | 0.552 | 0.6 | 0.601 | 0.583 | 0.635 |
| 행정안전부 | 0.484 | 0.552 | 0.581 | 0.509 | 0.548 |

* '14년 하반기에는 산업통상자원부와 국토교통부가 배포한 관련 보도자료 없음

제 4 절 사회기술시스템 척도 분석

1. 시기별 비교

단어 핵심어 분석은 각 단어가 얼마나 중요하게 사용되었는지 파악이 가능하지만, 각 단어의 연결 간의 맥락에서 어떤 의미로 사용되었는지를 아는 데에는 한계가 있다. 따라서 각 핵심어가 사회기술시스템적으로 어떤 맥락에서 사용되었는지 수치 도출을 시도하였다.

구체적으로 앞서 도출한 핵심어의 사회기술시스템 맥락을 위키피디아 검색량을 통해 도출한 상호정보량을 이용하였다. 상호정보량은 위키피디아 검색정보를 활용하여, 각 단어가 과학·기술, 산업·경제, 사회·제도 각 속성 중 어느 속성에 가까운 지를 수치화한다. 한편, 앞서 핵심어 분석에서 도출한 연결중심성은 해당 단어가 사회 연결망에서 상대적으로 중요한 척도를 나타내므로, 각 시기에서 해당 단어가 차지하는 중요성을 수치화한다.

본 연구는 각 시기와 기관을 비교를 위해 과학·기술, 산업·경제, 사회·제도별로 더한 값이 1이 되도록 표준화한 상호정보량을 연결중심성이 0.07이상인 단어의 연결중심성과 곱하고, 곱한 값들을 모두 더하였다. 본 연구는 이를 ‘사회기술시스템 척도’ 라고 하였으며, 이를 통해 각 시기별, 주요 중앙행정기관별 지능정보정책을 사회기술시스템 맥락에서 살펴볼 수 있다고 보았다. 도출 결과를 <표 9>에 제시하였다.

<표 9> 사회기술시스템 척도 (시기별)

| | 과학 기술 | 산업 경제 | 사회 제도 | 문서 수 | 문장 수 |
|---------|-------|-------|-------|-------|--------|
| 전체기간 | 0.346 | 0.356 | 0.298 | 1,143 | 12,755 |
| '14 상반기 | 0.354 | 0.372 | 0.274 | 26 | 1,413 |
| '14 하반기 | 0.416 | 0.329 | 0.255 | 16 | 299 |
| '15 상반기 | 0.368 | 0.374 | 0.259 | 52 | 545 |
| '15 하반기 | 0.376 | 0.362 | 0.262 | 71 | 640 |
| '16 상반기 | 0.396 | 0.346 | 0.258 | 101 | 981 |
| '16 하반기 | 0.391 | 0.359 | 0.250 | 158 | 1,758 |
| '17 상반기 | 0.354 | 0.378 | 0.268 | 258 | 2,734 |
| '17 하반기 | 0.378 | 0.362 | 0.260 | 241 | 2,795 |
| '18 상반기 | 0.370 | 0.351 | 0.278 | 185 | 2,758 |

전체시기를 통해 살펴보면, 중앙행정기관의 지능정보정책에 관한 보도자료는 산업·경제(0.356), 과학·기술(0.346), 사회·제도(0.298) 순으로 설명되고 있음을 확인할 수 있다.

산업·경제, 과학·기술 수치에 비해 사회·제도 수치가 상당히 낮게 나타남을 확인할 수 있다. 법·제도 정비, 일자리 대응 및 윤리 정립과 관련된 단어가 연결중심성 상위 단어에 속하지 못하였고, 이는 전반적으로 사회·제도 속성의 수치가 낮게 나타나는 원인이 되었다.

관계부처 합동(2016)과 4차산업혁명위원회(2017)는 지능정보기술로 인한 사회 변화를 대비하면서 과학·기술, 산업·경제, 사회·제도 전 분야의 변화에 맞추어 각 분야가 긴밀하게 연계된 종합 정책을 추진하겠다고 천명하였다. 하지만, 아직까지 정책의 관심은 기술적 측면에서 경쟁력 확보와 산업적 활용 측면에 집중되고 있으며, 지능정보기술의 도입으로 인한 사회 변화에 대응하는 교육 혁신, 일자리 변화 대응, 법제도 정비 및 윤리 정립과 같은 사회·제도적 주요 쟁점은 충분히 다루지 못하다고 해석할 수 있다.

한편, 시기별로 뚜렷한 변화의 경향성이 나타나지는 않다. 이 결과는 기술도입 초기에 과학·기술 맥락에 대한 설명이 상대적으로 많고, 기술이 실제 도입되고 적용되는 단계로 갈수록 사회·제도적인 환경 마련과 산업·경제적 활용에 초점을 맞춘 경우가 증가할 것이라는 예상과는 다른 결과이다. 2018년 상반기까지도 지능정보기술의 도입 단계가 도입 초기 단계이기 때문에 유의미한 변화가 나타나지 않은 것이라 본다.

2. 주요 기관별 비교

시기별 비교와 마찬가지로 각 기관별 연결 중심성(dgree centrality)이 높은 단어(0.07이상)를 뽑고, 상호정보량을 통해 도출한 사회기술시스템 척도로 해당 단어의 사회기술시스템적 맥락을 수치화 하였다.

구체적인 수치는 <표 10>에 정리하였으며, 다음과 같다. 과학기술정보통신부의 사회기술시스템 척도는 과학·기술(0.369), 산업·경제(0.362), 사회·제도(0.269) 순으로 나타났다. 산업통상자원부의 사회기술시스템 척도는 과학·기술(0.392), 산업·경제(0.371), 사회·제도(0.238) 순이다. 국토교통부의 사회기술시스템 척도는 과학·기술(0.382), 산업·경제(0.342), 사회·제도(0.276) 순이다. 행정안전부의 사회기술시스템 척도는 과학·기술(0.355), 산업·경제(0.342), 사회·제도(0.303) 순이다.

네 기관이 수치의 정도의 차가 있지만, 공통적으로 과학·기술, 산업·경제, 사회·제도 순으로 설명하고 있다고 할 수 있다. 네 기관 외에 다른 중앙행정기관을 모두 포함하여 도출한 값이 산업·경제, 과학·기술, 사회·제도 순인 것과 다른 결과를 나타내고 있다. 지능정보정책을 보다 많이 다루는 기관들이 그렇지 않은 기관에 비해 상대적으로 과학·기술 속성의 단어를 보다 많이 사용하고 있음을 알 수 있다. 지능정보정책을 많이 다루는 네 개 기관이 다른 기관들에 비해 과학·기술의 소개와 도입에 대한 내용을 주요하게 다루고 있으며, 그 외 기관은 기술 그 자체보다는 산업·경제적인 활용을 보다 주요하게 다루는 경향을 나타내고 있다고 할 수 있다.

<표 10> 사회기술시스템 척도 (기관별)

| | 과학 기술 | 산업 경제 | 사회 제도 | 문서 수 | 문장 수 |
|---------------|-------|--------|-------|-------|--------|
| 전체기관 | 0.346 | 0.356 | 0.298 | 1,143 | 12,755 |
| 과학기술 정보통신부 | 0.369 | 0.362 | 0.269 | 271 | 3,145 |
| 산업통상 자원부 | 0.392 | 0.371 | 0.238 | 122 | 1,191 |
| 국토교통부 | 0.382 | 0.3419 | 0.276 | 254 | 2,940 |
| 행정안전부 | 0.355 | 0.3420 | 0.303 | 112 | 1,417 |

한편, 네 기관의 상대적인 수치를 비교하면, 과학·기술 수치는 산업통상자원부, 과학기술정보통신부, 국토교통부, 행정안전부 순이며, 산업·경제 수치는 산업통상자원부, 과학기술정보통신부, 행정안전부, 국토교통부 순, 사회·제도 수치는 행정안전부, 국토교통부, 과학기술정보통신부, 산업통상자원부 순이다.

과학기술정보통신부와 산업통상자원부의 속성이 상대적으로 비슷하고, 국토교통부와 행정안전부의 속성이 상대적으로 비슷하다. 이는 앞서 지니계수를 활용한 단어분포의 집중도를 도출했을 때와 비슷하게 묶인다는 점에서 흥미로운 결과라고 할 수 있다.

과학기술 정책 전반에 대한 주무부서인 과학기술정보통신부와 산업·공업 정책을 다루는 산업통상자원부가 과학·기술 관련 설명이 많고, 전자정부와 행정혁신을 다루는 기관인 행정안전부와 국토·교통 분야와 관련한 특정 산업 정책을 담당하는 국토교통부는 상대적으로 사회·제도 관련 설명을 많이 하고 있다고 할 수 있다.

네 기관 공통적으로 사회·제도적 맥락으로 분류된 단어가 적은 편이며, 행정안전부가 사회·제도적 단어의 비중이 상대적으로 큰 편이다.

전반적으로 사회·제도적인 맥락에서 사용된 단어가 적으며, 이는 사회·제도적 측면에서 구체적인 계획보다는 추상적인 계획 위주로

서술된 탓으로 본다. 과학·기술, 산업·경제적인 맥락에서도 세부 내용을 살펴보았을 때, 추상적이거나 포괄적인 단어로 구성된 경우가 많아 사회기술시스템 접근에서 볼 때, 바람직하다고 평가하기는 어렵다. 아직까지 지능정보기술 도입의 초기 단계이며, 신사업 활성화가 중요하기 때문에 각 중앙 행정기관별로도 기술의 도입과 산업계에서의 활용에 초점을 맞춘 측면이 있다.

하지만 관계부처 합동(2016)과 4차산업혁명위원회(2017)는 지능정보기술로 인한 사회 변화를 대비하면서 과학·기술, 산업·경제, 사회·제도 전 분야의 변화에 맞추어 각 분야가 긴밀하게 연계된 종합 정책을 추진하겠다고 천명하였으며, 기술 도입 초기 단계부터 미래사회 변화에 선제적으로 대응하는 모습이 필요하다. 실제 주요 행정기관별로도 구체적인 사회·제도 관련 이슈가 상당하다. 예를 들어 과학기술정보통신부는 지능정보정책의 총괄부처로서 국가적 추진방향에 대한 법·제도 정비, 국토교통부는 자율주행차, 드론 상용화와 관련한 제도 개선, 윤리적 기준 마련, 행정안전부는 공공기관 활용을 위한 제도 정비, 산업통상자원부는 신산업 성장 동력 기반 마련을 위한 여건 개선 등과 관련 단어의 활용이 기대와 다르게 부족하다고 볼 수 있다.

다시 말해서 각 주요 행정기관은 지능정보 기술의 도입으로 인한 사회 변화에 대응하는 교육 혁신, 일자리 변화 대응, 법제도 정비 및 윤리 정립과 같은 사회·제도적 주요 쟁점을 충분히 다루지 못하고 있다. 즉, 각 주요 중앙행정기관의 정책의 관심은 아직까지 기술적 측면의 경쟁력 확보와 산업적 활용 측면에만 집중되고 있다.

네 기관의 시기별 사회기술시스템 척도의 흐름은 다음 <표 11>, <표 12>, <표 13>, <표 14>와 같다.

<표 11> 사회기술시스템 척도 (과학기술정보통신부)

| | 과학 기술 | 산업 경제 | 사회 제도 | 문서 수 | 문장 수 |
|---------|-------|-------|-------|------|-------|
| 전체기간 | 0.369 | 0.362 | 0.269 | 271 | 3,145 |
| '14 상반기 | 0.317 | 0.400 | 0.283 | 9 | 90 |
| '14 하반기 | 0.383 | 0.348 | 0.269 | 16 | 161 |
| '15 상반기 | 0.360 | 0.377 | 0.264 | 28 | 316 |
| '15 하반기 | 0.331 | 0.396 | 0.274 | 24 | 244 |
| '16 상반기 | 0.411 | 0.355 | 0.234 | 20 | 194 |
| '16 하반기 | 0.433 | 0.322 | 0.245 | 26 | 308 |
| '17 상반기 | 0.365 | 0.378 | 0.257 | 27 | 271 |
| '17 하반기 | 0.373 | 0.360 | 0.266 | 77 | 963 |
| '18 상반기 | 0.397 | 0.344 | 0.258 | 44 | 598 |

<표 12> 사회기술시스템 척도 (산업통상자원부)

| | 과학 기술 | 산업 경제 | 사회 제도 | 문서 수 | 문장 수 |
|---------|-------|-------|-------|------|-------|
| 전체기간 | 0.392 | 0.371 | 0.238 | 122 | 1,191 |
| '14 상반기 | 0.411 | 0.378 | 0.211 | 2 | 11 |
| '14 하반기 | 0.000 | 0.000 | 0.000 | — | — |
| '15 상반기 | 0.347 | 0.389 | 0.264 | 3 | 13 |
| '15 하반기 | 0.360 | 0.359 | 0.281 | 5 | 15 |
| '16 상반기 | 0.389 | 0.370 | 0.241 | 5 | 14 |
| '16 하반기 | 0.366 | 0.368 | 0.265 | 13 | 135 |
| '17 상반기 | 0.392 | 0.376 | 0.232 | 46 | 376 |
| '17 하반기 | 0.383 | 0.383 | 0.234 | 25 | 319 |
| '18 상반기 | 0.380 | 0.371 | 0.249 | 23 | 308 |

<표 13> 사회기술시스템 척도 (국토교통부)

| | 과학 기술 | 산업 경제 | 사회 제도 | 문서 수 | 문장 수 |
|---------|-------|-------|-------|------|-------|
| 전체기간 | 0.382 | 0.342 | 0.276 | 254 | 2,940 |
| '14 상반기 | 0.296 | 0.360 | 0.344 | 1 | 14 |
| '14 하반기 | 0.000 | 0.000 | 0.000 | — | — |
| '15 상반기 | 0.352 | 0.382 | 0.265 | 6 | 44 |
| '15 하반기 | 0.360 | 0.349 | 0.292 | 18 | 140 |
| '16 상반기 | 0.399 | 0.325 | 0.276 | 34 | 327 |
| '16 하반기 | 0.395 | 0.345 | 0.260 | 40 | 444 |
| '17 상반기 | 0.418 | 0.345 | 0.237 | 53 | 577 |
| '17 하반기 | 0.412 | 0.350 | 0.238 | 44 | 552 |
| '18 상반기 | 0.382 | 0.342 | 0.276 | 58 | 842 |

<표 14> 사회기술시스템 척도 (행정안전부)

| | 과학 기술 | 산업 경제 | 사회 제도 | 문서 수 | 문장 수 |
|---------|-------|-------|-------|------|-------|
| 전체기간 | 0.355 | 0.342 | 0.303 | 112 | 1,417 |
| '14 상반기 | 0.316 | 0.357 | 0.327 | 3 | 37 |
| '14 하반기 | 0.340 | 0.325 | 0.335 | 2 | 24 |
| '15 상반기 | 0.333 | 0.347 | 0.320 | 6 | 60 |
| '15 하반기 | 0.359 | 0.342 | 0.299 | 12 | 132 |
| '16 상반기 | 0.335 | 0.364 | 0.301 | 9 | 106 |
| '16 하반기 | 0.345 | 0.337 | 0.318 | 18 | 229 |
| '17 상반기 | 0.334 | 0.357 | 0.309 | 30 | 425 |
| '17 하반기 | 0.366 | 0.373 | 0.262 | 14 | 156 |
| '18 상반기 | 0.369 | 0.332 | 0.299 | 18 | 248 |

제 5 장 결론

제 1 절 연구결과 요약

본 연구는 지능정보기술 관련 정부 기관이 인식하고 있는 현재의 모습에 대한 분석이 필요하다는 문제의식 하에 논의를 진행하였다. 중앙행정기관 지능정보정책의 전개 양상과 특징은 무엇인가라는 연구문제를 가지고, 지능정보정책에 관한 보도자료를 분석하였다. 방법론으로는 텍스트를 알고리즘에 의해 분석하는 텍스트 마이닝(text mining) 기법을 사용하여 분석하였다.

분석 결과는 다음과 같다. 첫째, 중앙행정기관의 지능정보 정책의 전반적인 현황과 관련하여, 지능정보와 관련되는 키워드는 빅데이터, 드론, 4차 산업혁명, 클라우드컴퓨팅, 자율주행, 스마트시티, 인공지능 순으로 많이 언급되었다, 중앙행정기관별로는 과학기술정보통신부, 국토교통부, 산업통상자원부, 행정안전부 순서로 관련 보도자료를 많이 제공하였다. 또한, 분석기간을 반기(半期)별로 나누어 살펴볼 때, 2014년 상반기부터 2017년 상반기까지 관련 보도자료가 매 반기별로 증가하였으나, 2017년 하반기와 2018년 상반기는 2017년 상반기에 비해 관련 보도자료가 줄어들었음을 확인할 수 있다.

둘째, 중앙행정기관의 보도자료에 나타난 핵심어의 흐름을 시기별로 살펴보면 각 시기별 지능정보기술 활용의 초점이 다름을 알 수 있다. 분석 대상 초기 시기(2014년~2015년)의 경우, ‘업무와 서비스를 지원’ 하는 측면에 초점이 있었으나, 시간이 지나면서 ‘산업, 기업에 활용’ 하는 측면이 점차 강조되었다. 또, 각 시기별 자주 언급되는 지능정보기술이 다름을 알 수 있는데, 이전에 연결중심성 상위 단어로

등장하지 않던 ‘드론, 4차 산업혁명, 스마트 시티, 인공지능’ 이 2016년 이후 등장하였고, 이전에 연결중심성 상위에 속하던 ‘클라우드 컴퓨팅, 빅데이터, 사물인터넷’의 비중은 상대적으로 줄어들었다.

셋째, 주요 기관별 핵심어 분석을 통해 비교하면 행정안전부와 나머지 세 개 기관(과학기술정보통신부, 산업통상자원부, 국토교통부)의 추출된 단어 특성이 구분됨을 확인할 수 있다. 이는 행정안전부가 지원·통제 기능을 수행하는 참모(staff) 기관의 성격을 지니고, 과학기술정보통신부, 국토교통부, 산업통상자원부가 산업 진흥 등의 역할을 수행하는 경제 부처의 성격을 지닌다는 점이 핵심어 분석에서도 나타났다고 볼 수 있다.

넷째, 소득분포 정도를 측정하는 지니계수를 단어분포에 적용하여 도출한 단어분포의 집중도를 시기별로 살펴보면, 시간이 지날수록 단어분포의 집중도가 커지는 경향이 있음을 알 수 있다. 즉, 핵심단어가 지능정보정책을 설명하는 비율이 더 커지고 있음을 확인할 수 있다.

다섯째, 단어분포의 집중도를 각 기관별로 비교해보면 과학기술정보통신부, 산업통상자원부, 행정안전부, 국토교통부 순으로 특정 단어가 집중되어있는 정도가 크다고 확인할 수 있다. 한편, 각 개별 기관별로도 정도의 차이는 있지만, 시간이 지날수록 단어분포의 집중도가 커지고 있다.

여섯째, 핵심어가 사회기술시스템 관점에서 어떤 맥락에서 사용되었는지를 해석하기 위해 연결중심성과 상호정보량을 이용하여 도출한 사회기술시스템 척도를 도출하여 살펴보면, 중앙행정기관의 지능정보정책은 산업·경제, 과학·기술, 사회·제도의 환경 순으로 설명이 가능하다. 한편, 시기 흐름에 따른 변화의 경향성은 특별히 발견되지 않는다.

일곱째, 주요 행정기관별 사회기술시스템 척도를 살펴보면, 네 기관이 수치의 정도의 차가 있지만, 공통적으로 과학·기술, 산업·경제, 사회·제도 순으로 지능정보정책을 설명하고 있다고 할 수 있다. 네 기관 외에 다른 중앙행정기관을 모두 포함하여 도출한 값이 산업·경제, 과학·기술, 사회·제도 순인 것과 다른 결과를 나타내고 있다. 또한, 네 기관 공통적으로 사회·제도적 맥락으로 분류된 단어가 적은 편이나 행정안전부의 사회·제도적 단어 비중이 상대적으로 큰 편이다.

제 2 절 연구의 함의와 제언

본 연구는 대량의 비정형 텍스트를 알고리즘을 이용하여 해석함으로써 정부가 수행하고 있는 정책에 관한 현재의 모습을 보다 객관적으로 분석할 수 있는 가능성을 제시한다는 의미가 있다. 기존의 정성적 방법에 의한 내용분석은 연구자의 주관적 관점에 따라 요약과 해석이 자의적일 수 있으며, 무엇보다 텍스트의 분량이 방대할 경우 적용이 어렵다. 텍스트 마이닝을 통한 정부 자료의 분석은 이와 같은 한계를 보완할 수 있는 가능성을 제시한다. 단어분포의 집중도, 사회기술시스템 척도 등의 도출을 통해서도 텍스트 마이닝을 통해 정부 자료를 특정한 시각에서 해석할 수 있도록 조작적 정의가 가능함을 보았다.

본 연구에서는 중앙행정기관의 지능정보정책은 과학·기술, 산업·경제, 사회·제도 전 분야의 변화에 맞추어 각 분야가 긴밀하게 연계된 종합 정책이기 보다는 아직까지 정책의 관심은 기술적 측면에서 경쟁력 확보와 산업적 활용 측면에 집중된 정책임을 확인하였다. 사회기술시스템 관점에서 지능정보기술 도입과 관련 정책의 안착이 성공적으로 이루어지기 위해서는 과학·기술의 도입과 산업·경제 측면에서의 활용뿐만 아니라, 사회 변화에 대응하는 교육 혁신, 일자리 변화 대응, 법제도 정비 및 윤리 정립과 같은 사회·제도적 주요 쟁점을 다루는 정책 마련이 필요하다.

제 3 절 연구의 한계

본 연구는 기존의 정성적 연구방법을 사용한 내용 분석으로 분석하기 어려운 연구대상을 컴퓨터 연산을 이용해 해결하는 방법론인 텍스트 마이닝을 사용하였다는 점에서 의미가 있지만 연구의 한계 또한 명확하다.

첫째, 정책 브리핑에 올라온 보도자료가 곧 해당 중앙행정기관의 대표성 있는 정책인가에 대한 의문이 있을 수 있다. 보도자료는 행사에 대한 안내가 많고, 과학기술적으로 전문성 있는 내용을 깊이 있게 담지 못한다. 장기적인 정책 방향보다는 단기 홍보 위주의 내용의 비중 또한 상당한 것이 사실이다.

둘째, 기존에 진행된 자료를 되돌아보며 분석하는 데 머물기 때문에 향후 예측 모델을 제시하지는 못한다. 본 연구는 비정형 데이터로 구성된 보도자료를 컴퓨터 연산을 통해 분석할 수 있음을 확인하는 탐색적인 목적에 그치는 한계가 있어 연구모형 설정이 엄밀하지 못하다.

셋째, 연결중심성을 통한 핵심어 제시와 같이 내용 분석의 객관화를 시도하였으나, 결국 추출된 단어의 의미를 해석하는 것은 연구자의 몫이므로 의도와 달리 완전히 객관화된 연구라고 하기는 어려울 수 있다.

넷째, 사회기술시스템 척도로 제시한 수치의 엄밀성에 대한 점이다. 기존 선행 연구의 연구자의 주관에 의존하는 방식을 연산을 통해 객관성을 부여하는 목적을 가지고 있으나, 위키피디아 검색 정보를 통해 도출한 상호정보량의 단어의 맥락이 정부 보도자료에서 사용된 단어의 맥락과 일치한다는 보장은 없기 때문에 정부 보도자료로 구성된

사전을 별도 구성하여 보완할 필요가 있다. 또, 상호정보량은 단어와 단어 간의 관계를 분석하는 데 주로 사용되는 정보이기 때문에 범주와 단어 간의 관계에 적용하는 것에 무리가 따를 수 있다. 단어의 속성을 분류할 수 있는 보다 엄밀한 수치로 보완할 필요가 있다.

다섯째, 문서에서 특정 단어가 얼마나 대표성을 띄는 지에 대한 지표로서 빈도보다 연결중심성이 더 적절하다고 보았으나, 단어 간 네트워크를 살펴볼 때, 연결중심성 외에도 중심성에 대한 다양한 관점이 존재한다. 보도자료의 핵심어 도출에서 연결중심성이 가장 대표성을 띄는 지표라고 단언할 수는 없다.

참 고 문 헌

- 4차산업혁명위원회. (2017). 4차 산업혁명 정책방향.
- 관계부처 합동. (2016). 지능정보사회 중장기 종합대책.
- 관계부처 합동. (2017). 혁신성장을 위한 사람중심의 4차 산업혁명 대응계획
- 김동욱. (2018). 정부조직설계. 파주: 문우사.
- 김병운. (2016). 인공지능 동향분석과 국가차원 정책제언. 정보화정책, 23(1), 74-93.
- 김선경. (2014). 빅데이터 시대에 개인화서비스와 프라이버시 패러독스의 고찰. 한국지적정보학회지, 16(2), 193-207.
- 김영산, 왕규호. (2013). 미시경제학 (제2판). 서울: 박영사.
- 김윤명. (2016). 지능정보사회에 대한 규범적 논의와 법정정책적 대응. 정보화정책, 23(4), 24-37.
- 명승환, 허철준. (2012). 스마트사회 전환에 따른 Gov3.0 기반의 전자정부 개념과 패러다임 변화. 한국정책학회 춘계학술발표논문집, 2012, 325-341.
- 박주섭, 홍순구. (2016). 텍스트마이닝 기법을 활용한 부산시 지역 혁신정책 동향분석. 지방정부연구, 20(1), 1-20.
- 백영민. (2017). R을 이용한 텍스트 마이닝. 서울: 한울아카데미.
- 서병조, 신선영. (2017). 토픽모델링을 활용한 한국의 플랫폼 정부 연구동향 분석. 정보화정책, 24(3), 3-26.
- 성욱준, 황성수. (2017). 지능정보시대의 전망과 정책대응 방향 모색. 정보화정책, 24(2), 3-19.
- 성지은, 송위진. (2010). 탈(脫)추격형 혁신과 통합적 혁신정책. 과학기술학연구, 10(2), 1-36.
- 성지은, 정병걸, 송위진. (2012). 지속가능한 사회기술시스템으로의

- 전환과 백캐스팅. 과학기술학연구, 12(2), 81-116.
- 송위진. (2012). 사회·기술시스템론과 정책적 의의. 과학기술정책연구원 Issues & Policy, (60), 1-13.
- 송위진. (2013). 사회·기술시스템론과 과학기술혁신정책. 기술혁신학회지, 16(1), 156-175.
- 심주영. (2017). 용산미군기지 공원화 과정의 도시담론 분석. 한국도시설계학회지 도시설계, 18(5), 37-52.
- 어규철. (2011). 개도국 ICT인력의 효과적 양성을 위한 프로그램 개발협력 모델 연구. 연구보고서, 한국국제협력단.
- 이은미. (2018). Long Tail Phenomenon in Digital Government. 스마트시티와 지능형 정부를 위한 심포지엄. 서울대학교 행정대학원 지능정보사회 정책연구센터.
- 이은미, 김동욱, 고기동. (2016). 비정형 데이터 분석의 제도변화 적용에 관한 연구: 개인정보보호 제도를 중심으로. 한국정책과학학회보, 20(2), 217-240.
- 이재윤. (2003). 상호정보량의 정규화에 대한 연구. 한국문헌정보학회지, 37(4), 177-198.
- 이창길. (2010) 정권 초기의 가치지향과 정책우선순위 한국행정학보 44(3) 165-189.
- 이하영, 엄석진. (2018). 정부 지식관리의 성과와 영향요인에 관한 연구: 정보기술과 조직적 맥락 간 상호작용을 중심으로. 한국거버넌스학회보, 25(1), 257-296.
- 이현규. (2013). 사회기술적 접근방식을 통한 모바일 통신환경 분석. 한국산업정보학회논문지, 18(2), 59-69.
- 정보통신기술진흥센터. (2016). 주요 선진국의 제4차 산업혁명 정책동향.
- 정영미, 이재윤. (1998). 한국어 텍스트 내 용어연관성 분석을 위한 기초연구. 한국정보관리학회 학술대회 논문집, 243-246.

정준화, 김동욱. (2013). 스마트사회 개인정보 위험의 대상과 방식 연구. 한국지역정보학회지, 16(3), 113-136.

한국정보화진흥원. (2016). 새로운 기술, 새로운 세상 지능정보사회.

한재각, 조보영, 이진우. (2013). 걱정 ‘기술’에서 걱정 한 ‘사회기술 시스템’으로. 과학기술학연구, 13(2), 1-35.

황종성. (2016). 지능사회의 패러다임 변화 전망과 정책적 함의. 정보화정책, 23(2), 3-18.

Bostrom, R. P. and Heinen, J. S. (1997). "MIS Problems and Failures: A Socio-Technical Perspective PART II: THE Application of Socio-Technical Theory", MIS Quarterly, Vol.1, No.4. 11-29.

Chen, L. and Nath, R. (2008). "A socio-technical perspective of mobile work", Information Knowledge Systems Management, Vol.7, 41-60.

Cyram (2018). NetMinerVersion : 4.4.0.b Seoul: Cyram Inc.

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. Poetics, 41(6), 570-606.

Gentzkow, M., Kelly, B. T., & Taddy, M. (2017). Text as data.

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. Research policy, 31(8-9), 1257-1274.

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. Research

- policy, 33(6-7), 897-920.
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, 21(3), 267-297.
- Krippendorff, K. (2013). *Content analysis: An introduction to its methodology*. CA: Sage.
- Lyytinen, K., & Newman, M. (2008). Explaining information systems change: a punctuated socio-technical change model. *European Journal of Information Systems*, 17(6), 589-613.
- Orlikowski, W., and Iacono, C. (2001). Research commentary: desperately seeking the 'IT' in IT research - a call to theorizing the IT artifact, *Information Systems Research*, Co.12, No.2, 121-134.
- Reddick, C. G. (2012). *Public administration and information technology*. Burlington, MA: Jones & Bartlett Learning.
- Sawyer, S., Allen, J. P., & Lee, H. (2003). Broadband and mobile opportunities: a socio technical perspective. *Journal of Information Technology*, 18(2), 121-136.
- Schwab, K. (2016). *The Fourth Industrial Revolution*: World Economic Forum.

Abstract

An analysis of government
documents on intelligence
information policy: Focusing on
the press releases of the central
administrative agency

Choi, Hanbyul

Major in Public Administration

Graduate School of Public Administration

Seoul National University

A category of general purpose technology with applications to various fields, Intelligent information technology has extensive socio-economic effects—including those on the manufacturing, e-commerce, and finance sectors, as well as housing, transportation, and urban design. Considering the pervasiveness of intelligent information technology and its significance in contemporary society, this thesis examines the development and characteristics of the intelligence information policy adopted by the central administrative agency.

In order to do so, this thesis employed the text mining method to analyze text according to an algorithm. Analyzed documents are press releases from the central administrative agency regarding their intelligence information policy from January 1, 2014 to June 30, 2018. Consequently, this study analyzes a large amount of unstructured text using an algorithm, allowing for a more objective examination of the current state of government policies. Moreover, by utilizing text mining, this study overcomes the limitations associated with using qualitative methods for context analysis—especially for large amounts of text.

The results of the analysis are as follows. First, the focus of the flow of key words in the press releases on the agency's intelligence information policy differs according to the timing and institutional use. Second, the concentration of the keyword word distribution used in intelligence information policy press releases increases with time. Third, from the perspective of social technology systems, press releases on intelligent information policy are predominantly explained in terms of industry, economy, science and technology. In contrast, only a small proportion of terms are associated with the social and institutional environment.

The intelligence information policy of the central administrative agency is aimed at a comprehensive policy in which each field is closely aligned with changes in science, technology, industry, the economy, society, and the institution itself.

Nonetheless, policy is predominantly geared toward industrial utilization and remains focused on securing technological competitiveness. From the perspective of social technology systems, however, the introduction of technology and utilization of the industry are not enough to ensure the successful

implementation of intelligent information technology and related policies. Consequently, this thesis argues that it is imperative to develop a policy addressing social and institutional issues—such as educational innovation and response to job changes—as well as reform both the legal system and ethics in response to social change.

keywords : Intelligence information policy, Intelligence
information technology, Socio-technical system,
Text mining, Press release

Student Number : 2012-23756